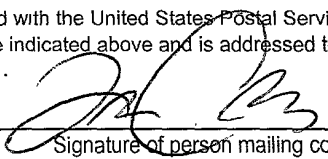


9-5-00

A

S9

09/01/00
 3675 U.S. PTO

Certificate of Mailing	
Date of Deposit <u>September 1, 2000</u>	Label Number: <u>EL509050555US</u>
I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as "Express Mail Post Office to Addressee" with sufficient postage on the date indicated above and is addressed to: BOX PATENT APPLICATION, Commissioner for Patents, Washington, D.C. 20231.	
<u>Luis A. Cruz</u> Printed name of person mailing correspondence	 Signature of person mailing correspondence

JC905 U.S. PTO
 09/654323
 09/01/00

UTILITY PATENT APPLICATION TRANSMITTAL UNDER 37 CFR §1.53(b)	
Attorney Docket Number	50110/004002
Applicant	Michael R. Hayden, Angela R. Brooks-Wilson, Simon N. Pimstone, and Susanne M. Clee
Title	COMPOSITIONS AND METHODS FOR MODULATING HDL CHOLESTEROL AND TRIGLYCERIDE LEVELS
PRIORITY INFORMATION:	
This application is a continuation-in-part of and claims priority from United States patent application 60/124,702, filed March 15, 1999; 60/138,048, filed June 8, 1999; 60/139,600, filed June 17, 1999, 60/151,977, filed September 1, 1999; 09/526,193, filed March 15, 2000; and 60/213,958 filed June 23, 2000.	
APPLICATION ELEMENTS:	
Cover sheet	1 page
Specification	105 pages
Claims	19 pages
Abstract	1 page
Drawing	86 sheets
Combined Declaration and POA, which is: <input checked="" type="checkbox"/> Unsigned; <input type="checkbox"/> Newly signed for this application; <input type="checkbox"/> A copy from prior application [**SERIAL NUMBER**] and the entire disclosure of the prior application is considered as being part of the disclosure of this new application and is hereby incorporated by reference therein.	3 pages
Statement Deleting Inventors	0 pages
Sequence Statement	2 pages
Sequence Listing on Paper	108 pages
Sequence Listing on Diskette	1 disk

Small Entity Statement, which is: <input checked="" type="checkbox"/> Unsigned; <input type="checkbox"/> Newly signed for this application; <input type="checkbox"/> A copy from prior application [**SERIAL NUMBER**] and such small entity status is still proper and desired.	2 pages
Preliminary Amendment	0 pages
IDS	0 pages
Form PTO 1449	0 pages
Cited References	0 references
Recordation Form Cover Sheet and Assignment	0 pages
Assignee's Statement	0 pages
English Translation	0 pages
Certified Copy of Priority Document	0 pages
Return Receipt Postcard	1
FILING FEES:	
Basic Filing Fee: \$345	\$345.00
Excess Claims Fee: 99 - 20 x \$9	\$711.00
Excess Independent Claims Fee: 45 - 3 = 42 x \$39	\$1638.00
Multiple Dependent Claims Fee: \$260/\$130	\$0
Total Fees:	\$2694.00
<input checked="" type="checkbox"/> Enclosed is a check for \$2694.00 to cover the total fees. <input type="checkbox"/> Charge [**AMOUNT**] to Deposit Account No. 03-2095 to cover the total fees. <input type="checkbox"/> The filing fee is not being paid at this time. <input checked="" type="checkbox"/> Please apply any other charges, or any credits, to Deposit Account No. 03-2095.	
CORRESPONDENCE ADDRESS:	
Paul T. Clark Reg. No. 30,162 Clark & Elbing LLP 176 Federal Street Boston, MA 02110 <div style="text-align: right;">Telephone: 617-428-0200 Facsimile: 617-428-7045</div>	
<div style="display: flex; justify-content: space-between;"> <div> <i>Susan M. Michael</i> Signature <i>Susan M. Michael Reg. No 42,285</i> </div> <div> <i>September 1, 2000</i> Date </div> </div>	

Applicant or Patentee : Michael R. Hayden et al.
Serial or Patent No. : Not yet assigned
Filed or Issued : September 1, 2000
Title : COMPOSITIONS AND METHODS FOR MODULATING HDL CHOLESTEROL AND TRIGLYCERIDE LEVELS

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
(37 CFR 1.9(f) and 1.27(d)) - NONPROFIT ORGANIZATION

I hereby declare that I am an official empowered to act on behalf of the nonprofit organization identified below:

Name of Organization: The University of British Columbia
Address of Organization: 2329 West Mall, Vancouver, British Columbia, V6T 1Z4 Canada
Type of Organization:

- ☒ University or Other Institution of Higher Education
☐ Tax Exempt under Internal Revenue Service Code (26 USC 501(a) and 501(c)(3))
☐ Nonprofit Scientific or Educational under Statute of State of the United States of America
 Name of State:
 Citation of Statute:
☐ Would Qualify as Tax Exempt under Internal Revenue Service Code (26 USC 501(a) and 501(c)(3)) If Located in the United States of America
☐ Would Qualify as Nonprofit Scientific or Educational under Statute of State of the United States of America If Located in the United States of America
 Name of State:
 Citation of Statute:

I hereby declare that the nonprofit organization identified above qualifies as a nonprofit organization as defined in 37 CFR 1.9(e) for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code with regard to the invention entitled COMPOSITIONS AND METHODS FOR MODULATING HDL CHOLESTEROL AND TRIGLYCERIDE LEVELS by inventors Michael R. Hayden, Angela R. Wilson, Simon N. Pimstone, and Susanne M. Clee described in

- ☒ the specification filed herewith.
☐ application serial No., [**SERIAL NUMBER**], filed [**FILING DATE**].
☐ patent no. [**PATENT NUMBER**], issued [**ISSUE DATE**].

I hereby declare that rights under contract or law have been conveyed to and remain with the nonprofit organization with regard to the above identified invention.

If the rights held by the nonprofit organization are not exclusive, each individual, concern or organization having rights to the invention is listed below* and no rights to the invention are held by any person, other than the inventor, who could not qualify as a small business concern under 37 CFR 1.9(c) or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

*NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

Full Name: Xenon Genetics, Inc.
Address: Suite #100, 2386 East Mall, Vancouver, British Columbia, CANADA V6T 1Z3
☐ INDIVIDUAL ☒ SMALL BUSINESS CONCERN ☐ NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

Name: Angus Livingstone

Title: Managing Director

Address: The University of British Columbia, 2329 West Mall, Vancouver, B.C. CANADA V6T 1Z4

Signature: _____ **Date:** _____

Applicant or Patentee : Michael R. Hayden et al.
 Serial or Patent No. : Not yet assigned
 Filed or Issued : September 1, 2000
 Title : COMPOSITIONS AND METHODS FOR MODULATING HDL CHOLESTEROL AND TRIGLYCERIDE LEVELS

**VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
 (37 CFR 1.9(f) and 1.27(c)) - SMALL BUSINESS CONCERN**

I hereby declare that I am

- ☐ the owner of the small business concern identified below:
☒ an official of the small business concern empowered to act on behalf of the concern identified below:

Name of Small Business Concern: Xenon Genetics, Inc.

Address of Small Business Concern: 2386 East Mall, Suite 100, Vancouver, B.C. V6T 1Z3 Canada

I hereby declare that the above identified small business concern qualifies as a small business concern as defined in 13 CFR 121.12, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees to the United States Patent and Trademark Office, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention, entitled COMPOSITIONS AND METHODS FOR MODULATING HDL CHOLESTEROL AND TRIGLYCERIDE LEVELS by inventors Michael R. Hayden, Angela R. Wilson, Simon N. Pimstone, and Susanne M. Clee described in

- ☒ the specification filed herewith.
☐ application serial no. [**SERIAL NUMBER**], filed [**FILING DATE**].
☐ patent no. [**PATENT NUMBER**], issued [**ISSUE DATE**].

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights to the invention is listed below and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d), or a nonprofit organization under 37 CFR 1.9(e). *NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

Assignee Name: The University of British Columbia

Assignee Address: 2329 West Mall, Vancouver, British Columbia, V6T 1Z4 CANADA

☐ INDIVIDUAL ☐ SMALL BUSINESS CONCERN ☒ NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent on which this verified statement is directed.

Name: Frank Holler

Title: Chief Executive Officer

Address: Xenon Genetics, Inc., 2386 East Mall, Suite 100, Vancouver, BC V6T 1Z3

Signature: _____ **Date:** _____

Certificate of Mailing

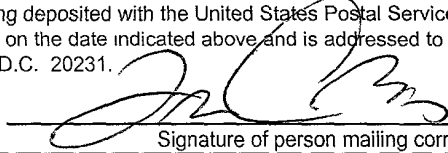
Date of Deposit September 1, 2000

Label Number: EL509050555US

I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as **"Express Mail Post Office to Addressee"** with sufficient postage on the date indicated above and is addressed to BOX PATENT APPLICATION, Commissioner for Patents, Washington, D.C. 20231.

Luis A. Cruz

Printed name of person mailing correspondence



Signature of person mailing correspondence

APPLICATION
FOR
UNITED STATES LETTERS PATENT

APPLICANT : Michael R. Hayden, Angela R. Brooks-Wilson, Simon N. Pimstone, and Susanne M. Clee
TITLE : COMPOSITIONS AND METHODS FOR MODULATING HDL CHOLESTEROL AND TRIGLYCERIDE LEVELS

COMPOSITIONS AND METHODS FOR MODULATING HDL
CHOLESTEROL AND TRIGLYCERIDE LEVELS

Cross Reference To Related Applications

This application is a continuation-in-part of U.S. Provisional Application No. 60/213958, filed June 23, 2000, and U.S. Utility Application No. 09/526,193, filed March 15, 2000, which claims priority from U.S. Provisional Application No. 60/124,702, filed March 15, 1999; U.S. Provisional Application No. 60/138,048, filed June 8, 1999; U.S. Provisional Application No. 60/139,600, filed June 17, 1999; and U.S. Provisional Application No. 60/151,977, filed September 1, 1999.

Background of the Invention

Low HDL cholesterol (HDL-C), or hypoalphalipoproteinemia, is a blood lipid abnormality which correlates with a high risk of cardiovascular disease (CVD), in particular coronary artery disease (CAD), but also cerebrovascular disease, coronary restenosis, and peripheral vascular disease. HDL-C levels are influenced by both environmental and genetic factors.

Epidemiological studies have consistently demonstrated that plasma HDL-C concentration is inversely related to the incidence of CAD. HDL-C levels are a strong graded and independent cardiovascular risk factor. Protective effects of an elevated HDL-C persist until 80 years of age. A low HDL-C is associated with an increased CAD risk even with normal (<5.2 mmol/l) total plasma cholesterol levels. Coronary disease risk is increased by 2% in men and 3% in women for every 1 mg/dL (0.026 mmol/l) reduction in HDL-C and in the majority of studies

2
this relationship is statistically significant even after adjustment for other lipid and non-lipid risk factors. Decreased HDL-C levels are the most common lipoprotein abnormality seen in patients with premature CAD. Four percent of patients with premature CAD have an isolated form of decreased HDL-C levels with no other
5 lipoprotein abnormalities while 25% have low HDL-C levels with accompanying hypertriglyceridemia.

Even in the face of other dyslipidemias or secondary factors, HDL-C levels are important predictors of CAD. In a cohort of diabetics, those with isolated low HDL-C had a 65% increased death rate compared to diabetics with normal HDL-C
10 levels (>0.9 mmol/l). Furthermore, it has been shown that even within high risk populations, such as those with familial hypercholesterolemia, HDL-C level is an important predictor of CAD. Low HDL-C levels thus constitute a major, independent, risk for CAD.

These findings have led to increased attention to HDL-C levels as a focus
15 for treatment, following the recommendations of the National Cholesterol Education Program. These guidelines suggest that HDL-C values below 0.9 mmol/l confer a significant risk for men and women. As such, nearly half of patients with CAD would have low HDL-C. It is therefore crucial that we obtain a better understanding of factors which contribute to this phenotype. In view of the
20 fact that pharmacological intervention of low HDL-C levels has so far proven unsatisfactory, it is also important to understand the factors that regulate these levels in the circulation as this understanding may reveal new therapeutic targets.

Absolute levels of HDL-C may not always predict risk of CAD. In the case of CETP deficiency, individuals display an increased risk of developing CAD,
25 despite increased HDL-C levels. What seems to be important in this case is the functional activity of the reverse cholesterol transport pathway, the process by

which intracellular cholesterol is trafficked out of the cell to acceptor proteins such as ApoAI or HDL. Other important genetic determinants of HDL-C levels, and its inverse relation with CAD, may reside in the processes leading to HDL formation and intracellular cholesterol trafficking and efflux. To date, this process is poorly understood, however, and clearly not all of the components of this pathway have been identified. Thus, defects preventing proper HDL-mediated cholesterol efflux may be important predictors of CAD. Therefore it is critical to identify and understand novel genes involved in the intracellular cholesterol trafficking and efflux pathways.

HDL particles are central to the process of reverse cholesterol transport and thus to the maintenance of tissue cholesterol homeostasis. This process has multiple steps which include the binding of HDL to cell surface components, the acquisition of cholesterol by passive absorption, the esterification of this cholesterol by LCAT and the subsequent transfer of esterified cholesterol by CETP, to VLDL and chylomicron remnants for liver uptake. Each of these steps is known to impact the plasma concentration of HDL.

Changes in genes for ApoAI-CIII, lipoprotein lipase, CETP, hepatic lipase, and LCAT all contribute to determination of HDL-C levels in humans. One rare form of genetic HDL deficiency is Tangier disease (TD), diagnosed in approximately 40 patients world-wide, and associated with almost complete absence of HDL-C levels (listed in OMIM as an autosomal recessive trait (OMIM 205400)). These patients have very low HDL-C and ApoAI levels, which have been ascribed to impairment of lipid transport and hypercatabolism of nascent HDL and ApoAI, due to a delayed acquisition of lipid and resulting failure of conversion to mature HDL. TD patients accumulate cholesterol esters in several tissues, resulting in characteristic features, such as enlarged yellow tonsils, corneal

opacity, hepatosplenomegaly, peripheral neuropathy, and cholesterol ester deposition in the rectal mucosa. Defective removal of cellular cholesterol and phospholipids by ApoAI as well as a marked deficiency in HDL mediated efflux of intracellular cholesterol has been demonstrated in TD fibroblasts. Even though this is a rare disorder, defining its molecular basis could identify pathways relevant for cholesterol regulation in the general population. The decreased availability of free cholesterol for efflux in the surface membranes of cells in Tangier Disease patients appears to be due to a defect in cellular lipid metabolism or trafficking. Approximately 45% of Tangier patients have signs of premature CAD, suggesting a strong link between decreased cholesterol efflux, low HDL-C and CAD. As increased cholesterol is observed in the rectal mucosa of persons with TD, the molecular mechanism responsible for TD may also regulate cholesterol adsorption from the gastrointestinal (GI) tract.

A more common form of genetic HDL deficiency occurs in patients who have low plasma HDL-C usually below the 5th percentile for age and sex (OMIM 10768), but an absence of clinical manifestations specific to Tangier disease (Marcil *et al.*, Arterioscler. Thromb. Vasc. Biol. 19:159-169, 1999; Marcil *et al.*, Arterioscler. Thromb. Vasc. Biol. 15:1015-1024, 1995). These patients have no obvious environmental factors associated with this lipid phenotype, and do not have severe hypertriglyceridemia nor have known causes of severe HDL deficiency (mutations in ApoAI, LCAT, or LPL deficiency) and are not diabetic. The pattern of inheritance of this condition is most consistent with a Mendelian dominant trait (OMIM 10768).

The development of drugs that regulate cholesterol metabolism has so far progressed slowly. Thus, there is a need for a better understanding of the genetic components of the cholesterol efflux pathway. Newly-discovered components can

then serve as targets for drugs.

Summary of the Invention

In a first aspect, the invention features a method for treating a patient
5 diagnosed as having a lower than normal HDL-cholesterol level or a higher than
normal triglyceride level. The method includes administering to the patient a
compound that modulates LXR-mediated transcriptional activity. Preferably, the
compound is administered to the patient with a pharmaceutically acceptable
carrier. The compound may be selected, for example, from the group consisting of
10 24-(S),25-epoxycholesterol; 24(S)-hydroxycholesterol; 22-(R)-hydroxycholesterol;
24(R),25-epoxycholesterol; 22(R)-hydroxy-24(S),25-epoxycholesterol;
22(S)-hydroxy-24(R),25-epoxycholesterol; 24-(S),25-iminocholesterol;
methyl-38-hydroxychololate; N,N-dimethyl-3 β -hydroxycholamide;
24(R)-hydroxycholesterol; 22(S)-hydroxycholesterol;
15 22(R),24(S)-dihydroxycholesterol; 25-hydroxycholesterol;
22(R)-hydroxycholesterol; 22(S)-hydroxycholesterol;
24(S),25-dihydroxycholesterol; 24(R),25-dihydroxycholesterol;
24,25-dehydrocholesterol; 25-epoxy-22(R)-hydroxycholesterol;
20(S)-hydroxycholesterol; (20R,22R)-cholest-5-ene-3 β ,20,22-triol;
20 4,4-dimethyl-5- α -cholesta-8,14,24-trien-3- β -ol;
7 α -hydroxy-24(S),25-epoxycholesterol; 7 β -hydroxy-24(S),25-epoxycholesterol;
7-oxo-24(S),25-epoxycholesterol; 7 α -hydroxycholesterol; 7-oxocholesterol;
and desmosterol. In one preferred embodiment, the compound is an oxysterol.

In a second aspect, the invention features another method for treating a
25 patient diagnosed as having a lower than normal HDL-cholesterol level or a higher
than normal triglyceride level. This method includes administering to the patient a

compound that modulates RXR-mediated transcriptional activity. RXR-modulating compounds include hetero ethylene derivatives; tricyclic retinoids; trienoic retinoids; benzocycloalkenyl-alka:di- or trienoic acid derivatives; bicyclic-aromatic compounds and their derivatives; bicycylmethyl-aryl acid derivatives; phenyl-methyl heterocyclic compounds; tetrahydro-naphthyl compounds; arylthio-tetrahydro-naphthalene derivatives and heterocyclic analogues; 2,4-pentadienoic acid derivatives; tetralin-based compounds; nonatetraenoic acid derivatives; SR11237; dexamethasone; hydroxy, epoxy, and carboxy derivatives of methoprene; bicyclic benzyl, pyridinyl, thiophene, furanyl, and pyrrole derivatives; benzofuran-acrylic acid derivatives; aryl-substituted and aryl and (3-oxo-1-propenyl)-substituted benzopyran, benzothiopyran, 1,2-dihydroquinoline, and 5,6-dihydronaphthalene derivatives; vitamin D3 (1,25-dihydroxyvitamin D3) and analogs; 24-hydroxylase inhibitor; mono-or polyenic carboxylic acid derivatives ; tetrahydroquinolin-2-one-6 or 7-yl and related derivatives; tetrahydronaphthalene; oxyiminoalkanoic acid derivatives; LG 100268; and LGD 1069.

In a third aspect, the invention features a method for determining whether a candidate compound modulates ABC1 expression by performing the steps of: (a) providing a nucleic acid molecule that includes an ABC1 regulatory region or promoter operably linked to a reporter gene; (b) contacting the nucleic acid molecule with the candidate compound; and (c) measuring expression of the reporter gene, wherein altered reporter gene expression, relative to a control not contacted with the compound, indicates that the candidate compound modulates ABC1 expression. In various preferred embodiments, the regulatory region includes 50 or more consecutive amino acids selected from nucleotides 5854 to 6694, 7756 to 8318, 10479 to 10825, 15214 to 16068, 21636 to 22111, 27898 to

28721, 32951 to 33743, 36065 to 36847, 39730 to 40577, 4543 to 5287, or 45081 to 55639 of SEQ ID NO: 1. In other preferred embodiments, the regulatory region 50 or more consecutive amino acids selected from nucleotides 1 to 28,707 or 29,011 to 53,228 of SEQ ID NO: 1. Preferably, the regulatory region includes a binding site for a transcription factor selected from a group consisting of LXRs, RXRs, RORs, SREBPs, and PPARs.

In a fourth aspect, the invention features a method for determining whether a person has an altered risk for developing cardiovascular disease. This method includes examining the person's ABC1 gene for polymorphisms or mutations. The presence of a polymorphism or mutation associated with cardiovascular disease indicates the person has an altered risk for developing cardiovascular disease.

In a related aspect, the invention features a method for predicting a person's response to a drug by determining whether the person has a polymorphism in an ABC1 gene that alters the person's response to the drug. Preferred polymorphisms are depicted in Fig. 4. In preferred embodiments of the fifth and sixth aspects, the polymorphism is in the 5' regulatory region of ABC1.

In a sixth aspect, the invention features a substantially purified LXR response element comprising the nucleotide sequence AGATCANNNNAGGTCA, wherein each N is, independently, C, T, G, or A (SEQ ID NO: 231). Preferably, the LXR response element has the sequence AGATCACTTGAGGTCA (SEQ ID NO: 232). Even more preferably, the LXR response element consists essentially of the nucleotide sequence AGATCANNNNAGGTCA, wherein each N is, independently, C, T, G, or A (SEQ ID NO: 231).

In a seventh aspect, the invention features a substantially pure nucleic acid molecule that consists essentially of a region that is substantially identical to at least 50, 100, 150, 300, 500, 750, 1000, 2000, 3000, 4000, 5000 or all of the

consecutive nucleotides selected from nucleotides 5854 to 6694, 7756 to 8318, 10479 to 10825, 15214 to 16068, 21636 to 22111, 27898 to 28721, 32951 to 33743, 36065 to 36847, 39730 to 40577, 45081 to 55639, 4543 to 5287, 59188 to 60306, 60689 to 63548, 63574 to 65110, 65030 to 68312, 68605 to 73375, 73395 to 74692, 75586 to 77103, 74774 to 74920, 77519 to 87679, 87651 to 94160, 96916 to 97634, 94408 to 96595, 97807 to 98989, 100369 to 107171, 107179 to 107983, 108039 to 108998, 109222 to 118212, 118612 to 123911, 124586 to 138185, 137773 to 138393, 147497 to 148051, 158490 to 159118, 123718 to 125077, 137773 to 138912, , 139304 to 139699, 139351 to 146359, 146867 to 147637, 147733 to 149404, 149858 to 152699, 153064 to 153916, 153978 to 158516, 158719 to 160272, 160375 to 164458, 165279 to 169814, 164215 to 164592, 164786 to 165133, 165125 to 165429, 169882 to 170189, 170067 to 174018, 176845 to 178875, 179113 to 180606, and 181723 to 183284 of SEQ ID NO: 1. In a related aspect, the invention features a substantially pure nucleic acid molecule that has a region that is substantially identical to nucleotides 5854 to 6694, 7756 to 8318, 10479 to 10825, 15214 to 16068, 21636 to 22111, 27898 to 28721, 32951 to 33743, 36065 to 36847, 39730 to 40577, 45081 to 55639, 4543 to 5287, 59188 to 60306, 60689 to 63548, 63574 to 65110, 65030 to 68312, 68605 to 73375, 73395 to 74692, 75586 to 77103, 74774 to 74920, 77519 to 87679, 87651 to 94160, 96916 to 97634, 94408 to 96595, 97807 to 98989, 100369 to 107171, 107179 to 107983, 108039 to 108998, 109222 to 118212, 118612 to 123911, 124586 to 138185, 137773 to 138393, 147497 to 148051, 158490 to 159118, 123718 to 125077, 137773 to 138912, , 139304 to 139699, 139351 to 146359, 146867 to 147637, 147733 to 149404, 149858 to 152699, 153064 to 153916, 153978 to 158516, 158719 to 160272, 160375 to 164458, 165279 to 169814, 164215 to 164592, 164786 to 165133, 165125 to 165429, 169882 to 170189,

170067 to 174018, 176845 to 178875, 179113 to 180606, or 181723 to 183284 of SEQ ID NO: 1. Preferred nucleic acid molecules have a region that is substantially identical or identical to nucleotides 1 to 28,707 of SEQ ID NO: 1 or nucleotides 29,011 to 53,228 of SEQ ID NO: 1.

5 In an eighth aspect, the invention features a method of treating a human having a lower than normal HDL-cholesterol level, a higher than normal triglyceride level, or a cardiovascular disease, including administering to the human an ABC1 polypeptide, or a cholesterol- or triglyceride-regulating fragment thereof, or a nucleic acid molecule encoding an ABC1 polypeptide, or a
10 cholesterol- or triglyceride-regulating fragment thereof. In a preferred embodiment, the human has a low cholesterol or high triglyceride level relative to normal. Preferably, the ABC1 polypeptide is wild-type ABC1, or has a mutation that increases its stability or its biological activity. Preferably, the nucleic acid molecule is operably linked to a promoter and contained in an expression vector.
15 Preferred mutations include the R \Rightarrow K mutation at position 219 and the V \Rightarrow A mutation at position 399 of ABC1. A preferred biological activity is improved regulation of cholesterol transport.

In a ninth aspect, the invention features a method of treating or preventing a lower than normal HDL-cholesterol level, a higher than normal triglyceride level,
20 or a cardiovascular disease, including administering to an animal (e.g., a human) a compound that mimics the activity of wild-type ABC1, R219K ABC1, or V399A ABC1 or modulates the biological activity of ABC1.

One preferred cardiovascular disease that can be treated using the methods of the invention is coronary artery disease. Others include cerebrovascular disease
25 and peripheral vascular disease.

The discovery that the ABC1 gene and protein are involved in cholesterol

transport that affects serum HDL levels allows the ABC1 protein and gene to be used in a variety of diagnostic tests and assays for identification of HDL-increasing, triglyceride-lowering, or CVD-inhibiting drugs. In one family of such assays, the ability of domains of the ABC1 protein to bind ATP is utilized;
5 compounds that enhance this binding are potential HDL-increasing or triglyceride-lowering drugs. Similarly, the anion transport capabilities and membrane pore-forming functions in cell membranes can be used for drug screening.

In a tenth aspect, ABC1 expression can also serve as a diagnostic tool for a lower than normal HDL-cholesterol level, a higher than normal triglyceride level,
10 or CVD; determination of the genetic subtyping of the *ABC1* gene sequence can be used to subtype individuals or families with lower than normal HDL levels or higher than normal triglyceride levels to determine whether the lower than normal HDL or higher than normal triglyceride phenotype is related to ABC1 function. This diagnostic process can lead to the tailoring of drug treatments according to
15 patient genotype (referred to as pharmacogenomics), including prediction of the patient's response (e.g., increased or decreased efficacy or undesired side effects upon administration of a compound or drug).

Antibodies to an ABC1 polypeptide can be used both as therapeutics and diagnostics. Antibodies are produced by immunologically challenging a B-cell-
20 containing biological system, e.g., an animal such as a mouse, with an ABC1 polypeptide to stimulate production of anti-ABC1 protein by the B-cells, followed by isolation of the antibody from the biological system. Such antibodies can be used to measure ABC1 polypeptide in a biological sample such as serum, by contacting the sample with the antibody and then measuring immune complexes as
25 a measure of the ABC1 polypeptide in the sample. Antibodies to ABC1 can also be used as therapeutics for the modulation of ABC1 biological activity.

Thus, in an eleventh aspect, the invention features a purified antibody that specifically binds to ABC1. In one preferred embodiment, the antibody modulates cholesterol or triglyceride levels when administered to a mammal.

In a twelfth aspect, the invention features a method for determining whether candidate compound is useful for modulating cholesterol or triglyceride levels, the method including the steps of: (a) providing an ABC1 polypeptide; (b) contacting the polypeptide with the candidate compound; and (c) measuring binding of the ABC1 polypeptide, wherein binding of the ABC1 polypeptide indicates that the candidate compound is useful for modulating cholesterol or triglyceride levels.

In a thirteenth aspect, the invention features a method for determining whether a candidate compound is useful for the treatment of a lower than normal HDL-cholesterol level, a higher than normal triglyceride level, or a cardiovascular disease. The method includes (a) providing an ABC transporter (e.g., ABC1); (b) contacting the transporter with the candidate compound; and (c) measuring ABC transporter biological activity, wherein increased ABC transporter biological activity, relative to a transporter not contacted with the compound, indicates that the candidate compound is useful for the treatment of a lower than normal HDL-cholesterol level, a higher than normal triglyceride level, or a cardiovascular disease. Preferably the ABC transporter is in a cell or a cell free assay system.

In a fourteenth aspect, the invention features a method for determining whether candidate compound is useful for modulating cholesterol or triglyceride levels. The method includes (a) providing a nucleic acid molecule comprising an ABC transporter promoter operably linked to a reporter gene; (b) contacting the nucleic acid molecule with the candidate compound; and (c) measuring expression of the reporter gene, wherein increased expression of the reporter gene, relative to a nucleic acid molecule not contacted with the compound, indicates that the

candidate compound is useful for modulating cholesterol or triglyceride levels.

In a fifteenth aspect, the invention features a non-human mammal having a transgene comprising a nucleic acid molecule encoding a mutated ABC1 polypeptide. In one embodiment, the mutation is a dominant-negative mutation, such as the M \Rightarrow T mutation at position 1091 of ABC1.

In a sixteenth aspect, the invention features an expression vector, a cell, or a non-human mammal that includes an *ABC1* nucleic acid molecule of the present invention.

In a related aspect, the invention features a cell from a non-human mammal having a transgene that includes a nucleic acid molecule of the present invention.

In an eighteenth aspect, the invention features a method for determining whether a candidate compound decreases the inhibition of a dominant-negative ABC1 polypeptide. The method includes (a) providing a cell expressing a dominant-negative ABC1 polypeptide; (b) contacting the cell with the candidate compound; and (c) measuring ABC1 biological activity of the cell, wherein an increase in the ABC1 biological activity, relative to a cell not contacted with the compound, indicates that the candidate compound decreases the inhibition of a dominant-negative ABC1 polypeptide. A preferred dominant-negative ABC1 polypeptide is M1091T ABC1.

In a nineteenth aspect, the invention features a method of determining in a subject a propensity for a disease or condition selected from the group consisting of a lower than normal HDL level, a higher than normal triglyceride level, and a cardiovascular disease. This method involves determining the presence or absence of at least one ABC1 polymorphism in the polynucleotide sequence of an ABC1 regulatory region, promoter, or coding sequence or in the amino acid sequence of an ABC1 protein in a sample obtained from the subject, wherein the presence or

absence of the ABC1 polymorphism is indicative of a risk for the disease or condition. Preferably, the method also includes analyzing at least five ABC1 polymorphic sites in the polynucleotide sequence or the amino acid sequence.

In a twentieth aspect, the invention features a method for determining whether an ABC1 polymorphism is indicative of a risk in a subject for a disease or condition selected from the group consisting of a lower than normal HDL level, a higher than normal triglyceride level, and a cardiovascular disease. The method includes (a) determining whether the prevalence of the disease or condition in a first subject or set of subjects differs from the prevalence of the disease or condition in a second subject or set of subjects; (b) analyzing the polynucleotide sequence of an ABC1 regulatory region, promoter, or coding sequence or the amino acid sequence of an ABC1 protein in a sample obtained from the first subject or set of subjects and the second subject or set of subjects; and (c) determining whether at least one ABC1 polymorphism differs between the first subject or set of subjects and the second subject or set of subjects, wherein the presence or absence of the ABC1 polymorphism is correlated with the prevalence of the disease or condition, thereby determining whether the ABC1 polymorphism is indicative of the risk. Preferably, the method further includes analyzing at least five ABC1 polymorphic sites in the polynucleotide sequence of an ABC1 regulatory region, promoter, or coding sequence or in the amino acid sequence of ABC1.

In a twenty-first aspect, the invention provides an electronic database having a plurality of sequence records of ABC1 polymorphisms correlated to records of predisposition to or prevalence of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

In a twenty-second aspect, the invention features a method for selecting a preferred therapy for modulating ABC1 activity or expression in a subject. This method includes (a) determining the presence or absence of at least one ABC1 polymorphism in the polynucleotide sequence of an ABC1 regulatory region, promoter, or coding sequence or in the amino acid sequence of an ABC1 protein in a sample obtained from the subject, wherein the presence or absence of the ABC1 polymorphism is indicative of the safety or efficacy of at least one therapy for modulating ABC1 expression or activity; and (b) determining a preferred therapy for modulating ABC1 expression or activity in the subject. Preferably, the method further includes analyzing at least five ABC1 polymorphic sites in the polynucleotide sequence of an ABC1 regulatory region, promoter, or coding sequence or the amino acids sequence of ABC1.

In a twenty-third aspect, the invention provides a method for determining whether a candidate compound is useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. This method includes (a) providing an assay system having a measurable ABC1 biological activity; (b) contacting the assay system with the candidate compound; and (c) measuring ABC1 biological activity or ABC1 phosphorylation. Modulation of ABC1 biological activity or ABC1 phosphorylation in this assay system, relative to the ABC1 biological activity or ABC1 phosphorylation in a corresponding control assay system not contacted with the candidate compound, indicates that the candidate compound is useful for the treatment of the disease or condition. In preferred embodiments, the assay system is a cell based system or a cell free system. Preferably, the candidate compound modulates both ABC1 protein phosphorylation and ABC1 activity.

In a twenty-fourth aspect, the invention provides a method for identifying a compound to be tested for an ability to ameliorate a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. This method includes (a) contacting a subject or cell with a candidate compound; (b) measuring ABC1 expression, activity, or protein phosphorylation in the subject or cell. Altered ABC1 expression, activity, or protein phosphorylation in this subject or cell; relative to the ABC1 expression, activity, or protein phosphorylation in a corresponding control subject or cell not contacted with the candidate compound; identifies the candidate compound as a compound to be tested for an ability to ameliorate the disease or condition. Preferably, the candidate compound modulates both ABC1 protein phosphorylation and the ABC1 activity.

In a twenty-fifth aspect, the invention provides a method for determining whether a candidate compound is useful for modulating a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. The method includes (a) providing a cell expressing an ABC1 gene or a fragment thereof; (b) contacting the cell with the candidate compound; and (c) measuring ABC1 activity of the cell. Altered ABC1 activity in this cell, relative to the ABC1 activity in a corresponding control cell not contacted with the compound, indicates that the candidate compound is useful for modulating the disease or condition.

In a twenty-six aspect, the invention provides a method for determining whether a candidate compound is useful for modulating a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. This method includes (a) contacting a cell expressing an ABC1 protein with the candidate

compound; and (b) measuring the phosphorylation of the ABC1 protein. Altered ABC1 protein phosphorylation in this cell, relative to the ABC1 protein phosphorylation in a corresponding control cell not contacted with the candidate compound, indicates that the is useful for modulating the disease or condition.

5 In a twenty-seventh aspect, the invention provides a compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. The compound modulates ABC1 biological activity, and is identified by the steps of (a) providing an assay system having a measurable
10 ABC1 biological activity; (b) contacting the assay system with the compound; and (c) measuring ABC1 biological activity, wherein modulation of ABC1 biological activity, relative to the ABC1 biological activity in a corresponding control assay system not contacted with the compound, indicates that the compound is useful for the treatment of the disease or condition.

15 In a twenty-eighth aspect, the invention provides a compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. The compound induces a change in ABC1 biological activity that mimics the change in ABC1 biological activity induced by the R219K
20 ABC1 mutation.

In a twenty-ninth aspect, the invention provides a compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. The compound binds or interacts with residue R219 of
25 ABC1, thereby mimicking the change in ABC1 activity induced by the R219K ABC1 mutation.

In a thirtieth aspect, the invention provides a compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. The compound induces a change in ABC1 biological activity that mimics the change in ABC1 biological activity induced by the V339A ABC1 mutation.

In a thirty-first aspect, the invention provides a compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. The compound binds or interacts with residue V399 of ABC1, thereby mimicking the change in ABC1 activity induced by the V399A ABC1 mutation.

In a thirty-second aspect, the invention provides a compound that modulates ABC1 activity and binds or interacts with an amino acid of ABC1, wherein the amino acid is a residue selected from amino acids 119 to 319 of ABC1 (SEQ ID NO: 5) or amino acids 299 to 499 of ABC1 (SEQ ID NO: 5).

In a thirty-second aspect, the invention provides a method for determining whether a candidate compound is useful for the treatment a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. This method involves (a) providing an assay system having a measurable LXR biological activity; (b) contacting the assay system with the candidate compound; and (c) measuring LXR biological activity, wherein modulation of LXR biological activity, relative to the LXR biological activity in a corresponding control assay system not contacted with the candidate compound, indicates that the candidate compound is useful for the treatment of the disease or condition.

In a thirty-third aspect, the invention provides method for determining whether a candidate compound is useful for modulating ABC1 biological activity. This method involves (a) providing an assay system having a measurable LXR biological activity; (b) contacting the assay system with the candidate compound; and (c) measuring LXR biological activity, wherein modulation of LXR biological activity, relative to the LXR biological activity in a corresponding control assay system not contacted with the candidate compound, indicates that the candidate compound is useful for modulating ABC1 biological activity. Preferably, the LXR biological activity is modulation of ABC1 expression.

In a thirty-fourth aspect, the invention provides method for identifying a compound to be tested for an ability to modulate ABC1 biological activity. This method involves (a) contacting a subject or cell with a candidate compound; (b) assaying the activity of the LXR gene product in the subject or cell; wherein modulation of the activity, relative to the activity in a corresponding control subject or cell not contacted with the candidate compound, identifies the candidate compound as a compound to be tested for an ability to modulate the biological activity of ABC1.

In a thirty-fifth aspect, the invention provides the use of an LXR gene product in an assay to identify compounds useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

In a thirty-sixth aspect, the invention features the use of a compound that modulates the activity or expression of an LXR gene product for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a

cardiovascular disease.

In a thirty-seventh aspect, the invention provides a method for identifying a compound to be tested for an ability to treat a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. This method involves

5 (a) providing an assay system having a measurable LXR biological activity;

(b) contacting the assay system with the candidate compound; and (c) measuring LXR biological activity, wherein modulation of the LXR biological activity, relative to the LXR biological activity in a corresponding control assay system not

10 contacted with the candidate compound, identifies the candidate compound as a compound to be tested for an ability to treat the disease or condition.

In a thirty-eight aspect, the invention provides a method for screening an candidate LXR agonist for the ability to treat a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. This method involves

15 (a) contacting the a with the candidate LXR agonist; and (b) measuring cholesterol efflux activity of the cell, wherein an increase in the cholesterol efflux activity in the cell, relative to the cholesterol efflux in a corresponding control cell not contacted with the candidate LXR agonist, indicates that the candidate LXR

20 agonist is useful for treating the disease or condition.

In a thirty-ninth aspect, the invention provides a method for screening a candidate LXR modulating compound for the ability to treat a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease. This method

25 involves (a) contacting a cell with the candidate LXR modulating compound; and (b) measuring ABC1 biological activity of the cell; wherein an increase in ABC1

biological activity in the cell, relative to the ABC1 biological activity in a corresponding control cell not contacted with the LXR modulating compound, indicates that the LXR modulating compound is useful for treating the disease or condition.

5 In another aspect, the invention provides a method for determining whether a candidate compound is useful for modulating triglyceride levels. The method involves (a) providing a cell comprising an ABC1 polypeptide comprising amino acids 1 to 60 of SEQ ID NO: 5; (b) contacting the cell with the candidate compound; and (c) measuring the half-life of the ABC1 polypeptide,
10 wherein an increase in said half-life, relative to the half-life in a corresponding control cell not contacted with the compound, indicates that the candidate compound is useful for modulating triglyceride levels.

In a related aspect, the invention features method for determining whether a candidate compound mimics ABC1 biological activity. The method includes (a)
15 providing a cell that is not expressing an ABC1 polypeptide; (b) contacting the cell with the candidate compound; and (c) measuring ABC1 biological activity of the cell, wherein altered ABC1 biological activity, relative to a corresponding control cell not contacted with the compound, indicates that the candidate compound modulates ABC1 biological activity. Preferably, the cell has an *ABC1* null
20 mutation. In one preferred embodiment, the cell is in a mouse or a chicken (e.g., a WHAM chicken) in which its *ABC1* gene has been mutated.

In a preferred embodiment of the screening methods of the present invention, the cell is in an animal. The preferred biological activity is transport of cholesterol (e.g., HDL cholesterol or LDL cholesterol) or interleukin-1, or is
25 binding or hydrolysis of ATP by the ABC1 polypeptide. Preferably, the ABC1 polypeptide used in the screening methods includes amino acids 1-60 of SEQ ID

NO: 5. Alternatively, the ABC1 polypeptide can include a region encoded by a nucleotide sequence that hybridizes under high stringency conditions to nucleotides 75 to 254 of SEQ ID NO: 6. Preferably, the subject is a human. Preferably, the cell or assay system has an exogenously supplied copy of an LXRE selected from the group consisting of SEQ ID NO: 94, SEQ ID NO: 92, and the LXRE consensus motif at nucleotide -7670 of the 3' end of intron 1. For various methods of the invention, a preferred LXR biological activity is modulation of ABC1 expression. A preferred LXR gene product is an ABC1 nucleic acid molecule or protein.

It is also contemplated that additional sequence of the ABC1 regulatory regions may be determined by sequencing the rest of the 4I8, 31J20, 47O19, or 179G21 Research Genetics RPCI-11 BACs using the methods described herein. Substantially pure nucleic acids containing regions substantially identical to at least 50, 100, 150, 300, 500, 750, 1000, 2000, 3000, 4000, 5000 consecutive nucleotides of these regions may be used in the methods of the present invention.

By "polypeptide" is meant any chain of more than two amino acids, regardless of post-translational modification such as glycosylation or phosphorylation.

By "reporter gene" is meant any gene which encodes a product whose expression is detectable and/or quantifiable by physical, immunological, chemical, biochemical, or biological assays. A reporter gene product may, for example, have one of the following attributes, without restriction: a specific nucleic acid/chip hybridization pattern, fluorescence (e.g., green fluorescent protein), enzymatic activity (e.g., lacZ/ β -galactosidase, luciferase, chloramphenicol acetyltransferase), toxicity (e.g., ricin A), or an ability to be specifically bound by a second molecule (e.g., biotin or a detectably labeled antibody). It is understood that any engineered

variants of reporter genes, which are readily available to one skilled in the art, are also included, without restriction, in the foregoing definition.

By "operably linked" is meant that a gene and a regulatory sequence are connected in such a way as to permit expression of the gene product under the control of the regulatory sequence. A promoter may also be operably linked to a gene such that expression of the gene product is under control of the promoter.

By "regulatory region" is meant a region that, when operably linked to a promoter and a gene (e.g., a reporter gene), is capable of modulating the expression of the gene from the promoter. Regulatory regions include, for example, nuclear hormone transcription factor binding sites such as those described herein and may be found in intronic sequence.

By "promoter" is meant a minimal sequence sufficient to direct transcription of an operably-linked gene.

By "substantially identical" is meant a polypeptide or nucleic acid exhibiting at least 50%, preferably 85%, more preferably 90%, and most preferably 95% identity to a reference amino acid or nucleic acid sequence. For polypeptides, the length of comparison sequences will generally be at least 16 amino acids, preferably at least 20 amino acids, more preferably at least 25 amino acids, and most preferably 35 amino acids. For nucleic acids, the length of comparison sequences will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 110 nucleotides. One sequence may include additions or deletions (i.e., gaps) of 20% or less when compared to the second sequence. Optimal alignment of sequences may be conducted, for example, by the methods of Gish and States (Nature Genet. 3:266-272, 1993), Altschul *et al.* (J. Mol. Biol. 215:403-410, 1990), Madden *et al.* (Meth. Enzymol. 266:131-141, 1996), Altschul *et al.* (Nucleic Acids

Res. 25:3389-3402, 1997), or Zhang *et al* (Genome Res. 7:649-656, 1997).

Sequence identity is typically measured using sequence analysis software with the default parameters specified therein (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). This software program matches similar sequences by assigning degrees of homology to various substitutions, deletions, and other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine, valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

By “substantially pure nucleic acid” is meant nucleic acid that is free of the genes which, in the naturally-occurring genome of the organism from which the nucleic acid of the invention is derived, flank the nucleic acid. The term therefore includes, for example, a recombinant nucleic acid that is incorporated into a vector; into an autonomously replicating plasmid or virus; into the genomic nucleic acid of a prokaryote or a eukaryote cell; or that exists as a separate molecule (e.g., a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. It also includes a recombinant nucleic acid that is part of a hybrid gene encoding additional polypeptide sequence.

By “high stringency conditions” is meant hybridization in 2X SSC at 40°C with a DNA probe length of at least 40 nucleotides. For other definitions of high stringency conditions, see F. Ausubel *et al.*, *Current Protocols in Molecular Biology*, pp. 6.3.1-6.3.6, John Wiley & Sons, New York, NY, 1994, hereby incorporated by reference.

By “modulates” is meant increase or decrease. Preferably, a compound that

modulates LXR-mediated transcription, RXR-mediated transcription, ABC1 gene expression, HDL-C levels, or triglyceride levels does so by at least 5%, more preferably by at least 10%, and most preferably by at least 25% or even at least 50%.

5 By “purified antibody” is meant antibody which is at least 60%, by weight, free from proteins and naturally occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably 90%, and most preferably at least 99%, by weight, antibody. A purified antibody may be obtained, for example, by affinity chromatography using recombinantly-
10 produced protein or conserved motif peptides and standard techniques.

By “specifically binds” is meant an antibody that recognizes and binds to, for example, a human ABC1 polypeptide but does not substantially recognize and bind to other non-ABC1 molecules in a sample, e.g., a biological sample, that naturally includes protein. A preferred antibody binds to the ABC1 polypeptide
15 sequence of Fig. 2A (SEQ ID NO: 5).

By “polymorphism” is meant that a nucleotide or nucleotide region is characterized as occurring in several different forms. A “mutation” is a form of a polymorphism in which the expression level, stability, function, or biological activity of the encoded protein is substantially altered.

20 By “LXR” is meant nuclear receptors LXR α and LXR β . Preferred LXRs include human LXR α (GenBank accession no. Q13133) and human LXR β (GenBank accession no. P55055)(see Apfel *et al.*, Mol. Cell. Biol. 14:7025-7035, 1994; Willy *et al.*, Genes Dev. 9:1033-1045, 1995; and Song *et al.*, Proc. Natl. Acad. Sci. USA 91:10809-10813, 1995, each of which is hereby incorporated by reference).

25 By “RXR” is meant nuclear receptors RXR α , RXR β , and RXR γ . Preferred RXRs include human RXR α (GenBank accession no. Q13133), human RXR β

(GenBank accession no. S37781), and human RXR γ .(GenBank accession no. Q13133).

By “ABC transporter” or “ABC polypeptide” is meant any transporter that hydrolyzes ATP and transports a substance across a membrane. Preferably, an ABC transporter polypeptide includes an ATP Binding Cassette and a transmembrane region. Examples of ABC transporters include, but are not limited to, ABC1, ABC2, ABCR, and ABC8.

By “ABC1 polypeptide” is meant a polypeptide having substantial identity to an ABC1 polypeptide having the amino acid sequence of SEQ ID NO: 5.

By “ABC biological activity” or “ABC1 biological activity” is meant hydrolysis or binding of ATP, transport of a compound (e.g., cholesterol, interleukin-1) or ion across a membrane, or regulation of cholesterol or phospholipid levels (e.g., either by increasing or decreasing HDL-cholesterol or LDL-cholesterol levels).

The invention provides methods for treating patients having low HDL-C and/or higher than normal triglyceride levels by administering compounds that modulate ABC1 biological activity or expression. For example, the compounds may modulate the transcriptional activity of LXR/RXR heterodimers. Many compounds that modulate LXR transcriptional activity or RXR transcriptional activity are known in the art. Preferred compounds of the invention are oxysterols; additional compounds are described herein.

The invention also provides screening procedures for identifying therapeutic compounds (cholesterol-modulating, triglyceride-modulating, or anti-CVD pharmaceuticals) which can be used in human patients. Compounds that modulate ABC1 biological activity or expression are considered useful in the invention, as are compounds that modulate ABC concentration, protein stability,

regulated catabolism, or its ability to bind other proteins or factors. In general, the screening methods of the invention involve screening any number of compounds for therapeutically active agents by employing any number of *in vitro* or *in vivo* experimental systems. Exemplary methods useful for the identification of such compounds are detailed below.

The methods of the invention simplify the evaluation, identification and development of active agents for the treatment and prevention of low HDL, higher than normal triglyceride levels, and CVD. In general, the screening methods provide a facile means for selecting natural product extracts or compounds of interest from a large population which are further evaluated and condensed to a few active and selective materials. Constituents of this pool are then purified and evaluated in the methods of the invention to determine their HDL-raising, triglyceride-lowering, anti-CVD activities, or a combination thereof.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof.

Brief Description of the Drawings

Fig. 1 shows the genomic sequence of human ABC1, including exons 1-50 (SEQ ID NO: 1). Capital letters denote exonic sequence, lower case letters denote 5' regulatory sequence or intronic sequence. "Z" denotes any nucleotide or other, including no nucleotide. The numbering used herein for the nucleotides in SEQ ID No:1 assumes that no nucleotide is present at the positions denoted "z;" however, it would be readily apparent to one skilled in the art that the numbering of the nucleotides in these sequences would change if a nucleotide is present at some or all of the positions denoted "z." "K" denotes nucleotides G or T; "Y" denotes nucleotides C or T; "R" denotes nucleotides A or G; "M" denotes nucleotides A or

C; “S” denotes nucleotides C or G; “H” denotes nucleotides A, C, or T; “B” denotes nucleotides C, G, or T. Because of the identification of an ABC1 exon upstream of exon 0 that we previously disclosed, the numbering used herein to refer to ABC1 exons has been increased by one compared to our previously used numbering (U.S.S.N. 09/526,193; U.S.S.N. 60/124,702; U.S.S.N. 60/138,048; 5 U.S.S.N. 60/139,600; U.S.S.N. 60/151,977). For example, the previously described exon 0 is referred to herein as exon 1.

Fig. 2A is the amino acid sequence of the human ABC1 protein (SEQ ID NO: 5). Fig. 2B is the nucleotide sequence of the human *ABC1* cDNA (SEQ ID 10 NO: 6).

Fig. 3 is a summary of locations of consensus transcription factor binding sites in the human ABC1 5’ regulatory sequence. The abbreviations are as follows: PPRE=peroxisome proliferator-activated receptor. SREBP=steroid response element-binding protein site. ROR=RAR-related orphan receptor. The 15 numbering used herein for the locations of the transcription factor binding sites assumes that no nucleotide is present at the positions denoted “z” in SEQ ID NO: 1. For the polymorphisms in the promoter region, the numbering is based on the first base of the promoter as nucleotide number -1. For exon 1, the numbering is based on the first base of exon 1 as nucleotide number +1. For the 5’ 20 end intron 1, the numbering is based on the first position in intron 1 as +1. For the 3’ end of intron 1, the numbering is based on the first base ‘5 to the start of exon 2 as nucleotide number -1.

Fig. 4 is a table summarizing polymorphisms in the genomic ABC1 sequences.

25 Figs. 5A and 5B are bar graphs showing the percent of heterozygotes or unaffected family members with HDL (Fig. 5A) and triglycerides (TG) (Fig 5B)

within a given range of percentiles for age and sex, based on the LRC criteria (Heiss *et al.*, Circulation 62:IV-116-IV-136, 1980). A broad distribution of HDL levels was seen in the heterozygotes, extending up to the 31st percentile for age and sex. Although there is overlap in the distribution of triglycerides between heterozygotes and unaffected family members, a larger portion of heterozygotes have triglyceride levels greater than the 80th percentile for age and sex.

Fig. 6 is a table characterizing TD patients, ABC1 heterozygotes, and unaffected family members.

Fig. 7 is a table summarizing the incidence of CAD in ABC1 heterozygotes.

Fig. 8 is a graph showing the average HDL levels in heterozygotes for each mutation versus the efflux levels measured in a heterozygous carrier of each mutation. The HDL levels are expressed as the percentage of the mean HDL level in the unaffected members of that family. The efflux levels are highly correlated with the levels of HDL cholesterol and are associated with 82% of the variation in HDL cholesterol levels.

Fig. 9 is a table summarizing the HDL levels and presence or absence of CAD in ABC1 heterozygotes. In the R2144X and R909X ABC1 mutations, the codon encoding Arg2133 or Arg909 is mutated to a STOP codon resulting in truncation of the encoded protein. In the "Del E,D 1893,94" mutation, the codons encoding Glu1893 and Asp1894 are deleted. The "invs25 + 1G-->C" mutation converts the first nucleotide of intron 25 from a "G" to a "C," removing a splice site. For the "del C6825-->2145X" mutation, the deletion of C6825 in the nucleotide sequence is a frame-shift mutation that results in a STOP codon at the codon corresponding to amino acid 2145 of the encoded protein. For the "CTC6952-4TT-->2203X" mutation, "CTC" is replaced by "TT" in the nucleotide sequence, resulting in the conversion of the codon encoding amino acid 2203 to a

stop codon.

Fig. 10 is a table comparing the mean lipid levels in unaffected family members and ABC1 heterozygotes with either missense or severe mutations.

Fig. 11 is a schematic diagram of the ABC1 protein, illustrating the location of the mutations and the presence or absence of CAD in carriers of the mutations. The number (n) of heterozygotes who are 40 years or older and may have developed CAD are listed.

Figs. 12A and 12B are pedigrees of two FHA kindreds, FHA3 and FHA1, respectively (Marcil *et al.*, Lancet 354:1341-1346, 1999). Males are denoted by square symbols, females by circles. Individuals heterozygous for mutations are given half-shaded symbols, with the probands indicated by arrows. A diagonal line indicates a deceased individual. The youngest individuals have HDL cholesterol at higher percentile ranges than those in the older generations.

Fig. 13 is a bar graph showing the percentage of individuals less than 30 years of age and from 30 to less than 70 years of age with HDL cholesterol levels in a given percentile range. Younger individuals have a far broader distribution of HDL cholesterol levels, clearly indicating that the impact of ABC1 on HDL levels is influenced by age.

Fig. 14 is a table summarizing HDL and TG levels in different age groups for ABC1 heterozygotes and unaffected family members.

Figs. 15A and 15B are graphs showing the mean HDL level in heterozygous males (Fig. 15A) and females (Fig. 15B) in 10 year age groups (plotted at the half-way point) compared to the 10th percentile distribution in the LRC population (Heiss *et al.*, *supra*). Error bars represent the standard deviation of each mean. The number of individuals in each group is shown under each data point. Beyond the age of 30, mean HDL levels in heterozygotes fall much lower

than the 10th percentile distribution; in contrast, mean HDL cholesterol levels in the heterozygotes less than 30 years old more closely approximate the 10th percentile distribution.

Figs. 16A and 16B are graphs showing the mean HDL (Fig. 16A) and triglyceride levels (Fig. 16B) in heterozygotes and unaffected family members falling within each tertile of BMI. The tertiles of BMI correspond to the following values: (1) BMI <21.4; (2) 21.4<BMI<25.1; (3) BMI>25.1.

Fig. 17 is a table showing the oligonucleotides and reaction conditions used for RFLP screening of ABC1 polymorphisms.

Fig. 18 is a picture of a gel showing RFLP genotyping of the R219K variant. The 177 base pair PCR product is not digested for the A allele, whereas the B allele is digested producing fragments of 107 and 70 base pairs.

Fig. 19 is a table showing the allele frequencies of polymorphisms in the ABC1 gene.

Fig. 20 is a table comparing MSD, MOD, and frequency of coronary events in R219K ABC1 variant carriers compared to controls.

Fig. 21 is a graph showing the event-free survival curves for carriers (AB+BB) and non-carriers (AA) of the R219K ABC1 variant. Carriers of the variant have a 29% increased event-free survival over the two years of the trial, compared with non-carriers.

Fig. 22 is a table showing the baseline demographics and lipid levels in the Regression Growth Evaluation Statin Study (REGRESS) cohort by R219K ABC1 genotype.

Fig. 23 is a table showing the lipid levels and CAD above and below the median age in R219K ABC1 carriers and controls.

Fig. 24 is a bar graph showing the percent difference in HDL cholesterol

levels between those greater and less than the median age (56.7 years) for each R219K genotype.

Figs. 25A and 25B are graphs showing the correlations of HDL cholesterol (Fig. 25A) and efflux (Fig. 25B) with age, by R219K genotype.

5 Figs. 26A and 26B are graphs showing the change in MSD (Fig. 26A) and MOD (Fig. 26B) by median age in carriers (AB+BB) and non-carriers (AA) of the R219K ABC1 variant.

Fig. 27 is a table showing the ethnic distribution of the R219K ABC1 variant.

Detailed Description

10 We have previously discovered that the human ABC1 (also known as ABCA1) genomic region contains consensus binding sites for transcription factors such as LXRs, RXRs, PPARs, SREBPs, and RORs. In the present invention, we report the sequence of additional regions of the ABC1 regulatory region which also contain consensus binding sites for transcription factors. We also discovered that heterozygotes for ABC1 mutations have age-modulated decreases in HDL, increases in triglyceride levels, and significantly increased risk for CAD. Furthermore, this phenotype was highly correlated with efflux, clearly demonstrating that impairment of reverse cholesterol transport is associated with decreased plasma HDL cholesterol, increased triglyceride levels, and increased atherogenesis. Accordingly, the present invention features screening methods to identify therapies that increase ABC1 function, resulting in increased plasma HDL cholesterol, decreased triglyceride levels, protection against atherosclerosis, or a combination of these effects.

25 Genes play a significant role influencing HDL levels. Tangier disease (TD)

was the first reported genetic HDL deficiency. Until recently, the molecular basis for TD was unknown, but now mutations in *ABCI* have been identified in TD patients (described below). For example, we have identified two additional probands and their families, and confirmed linkage and refined the locus to a limited genomic region. Mutations in the *ABCI* gene accounting for all four alleles in these two families were detected. A more frequent cause of low HDL levels is a distinct disorder, familial HDL deficiency (FHA). On the basis of independent linkage, meiotic recombinants and disease associated haplotypes, FHA was localized to a small genomic region encompassing the *ABCI* gene. A mutation in a conserved residue in *ABCI* segregated with FHA. Antisense reduction of the *ABCI* transcript in fibroblasts was associated with a significant decrease in cholesterol efflux.

Cholesterol is normally assembled with intracellular lipids and secreted, but in TD the process is diverted and cholesterol is degraded in lysosomes. This disturbance in intracellular trafficking of cholesterol results in an increase in intracellular cholesterol ester accumulation associated with morphological changes of lysosomes and the Golgi apparatus and cholesteryl ester storage in histiocytes, Schwann cells, smooth muscle cells, mast cells and fibroblasts.

The clinical and biochemical heterogeneity in patients with TD has led to the possibility that genetic heterogeneity may also underlie this disorder.

Considering this, we initially performed linkage analysis on these two families of different ancestries (TD-1 is Dutch, TD-2 is British; Frohlich *et al.*, Clin. Invest. Med. 10:377-382, 1987) and confirmed that the genetic mutations underlying TD in these families were localized to the same 9q31 region, to which a large family with TD had been assigned (Rust *et al.*, Nature Genetics 20:96-98, 1998).

Detailed haplotype analysis, together with the construction of a physical map,

refined the localization of this gene. Mutations in the ABC1 gene were found in TD.

FHA is much more common than TD, although its precise frequency is not known. While TD has been described to date in only 40 families, we have identified more than 40 FHA families in the Netherlands and Quebec alone. After initial suggestions of linkage to 9q31, thirteen polymorphic markers spanning approximately 10 cM in this region were typed and demonstrated the highest LOD score at D9S277. Analysis of the homozygosity of markers in the TD-2 proband, who was expected to be homozygous for markers close to TD due to his parents' consanguinity, placed the TD gene distal to D9S127. Combined genetic data from TD and FHA families pointed to the same genomic segment spanning approximately 1,000 kb between D9S127 and D9S1866. The *ABC1* transporter gene was contained within the minimal genomic region. RT-PCR analysis in one family demonstrated a deletion of leucine at residue 693 ($\Delta 693$) in the first transmembrane domain of ABC1, which segregated with the phenotype of HDL deficiency in this family.

ABC1 is part of the ATP-binding cassette (ABC transporter) superfamily, which is involved in energy-dependent transport of a wide variety of substrates across membranes (Dean *et al.*, Curr. Opin. Gen. Dev. 5:779-785, 1995). These proteins have characteristic motifs conserved throughout evolution which distinguish this class of proteins from other ATP binding proteins. In humans these genes essentially encode two ATP binding segments and two transmembrane domains (Dean *et al.*, Curr. Opin. Gen. Dev. 5:779-785, 1995). We have now shown that the ABC1 transporter is crucial for intracellular cholesterol transport.

We have demonstrated that reduction of the *ABC1* transcript using oligonucleotide antisense approaches results in decreased efflux, clearly

demonstrating the link between alterations in this gene and its functional effects. TD and FHA now join the growing list of genetic diseases due to defects in the ABC group of proteins including cystic fibrosis (Zielenski, *et al.*, Annu. Rev. Genet. 29:777-807, 1995), adrenoleukodystrophy (Mosser *et al.*, Nature 361: 726-730, 1993), Zellweger syndrome (Gärtner *et al.*, Nat. Genet. 1:23, 1992), progressive familial intrahepatic cholestasis (Bull *et al.*, Nat. Genet. 18:219-224, 1998), and different eye disorders including Stargardt disease (Allikmets *et al.*, Nat. Genet. 15:236-246, 1997), autosomal recessive retinitis pigmentosa (Allikmets *et al.*, Science 277:1805-1807, 1997), and cone-rod dystrophy (Cremers *et al.*, Hum. Mol. Genet. 7:355-362, 1998).

Patients with TD have been distinguished from patients with FHA on the basis that Tangier disease was an autosomal recessive disorder (OMIM 20540) while FHA is inherited as an autosomal dominant trait (OMIM 10768). Furthermore, patients with TD have obvious evidence for intracellular cholesterol accumulation which is not seen in FHA patients. It is now evident that heterozygotes for TD do have reduced HDL levels and that the same mechanisms underlie the HDL deficiency and cholesterol efflux defects seen in heterozygotes for TD as well as FHA. Furthermore, the more severe phenotype in TD represents loss of function from both alleles of the *ABCI* gene.

ABC1 is activated by protein kinases, presumably via phosphorylation, which also provides one explanation for the essential role of activation of protein kinase C in promoting cholesterol efflux (Drobnick *et al.*, Arterioscler. Thromb. Vasc. Biol. 15: 1369-1377, 1995). Brefeldin, which inhibits trafficking between the endoplasmic reticulum and the Golgi, significantly inhibits cholesterol efflux, essentially reproducing the effect of mutations in ABC1, presumably through the inhibition of ABC1 biological activity. This finding has significance for the

understanding of mechanisms leading to premature atherosclerosis. TD homozygotes develop premature coronary artery disease, as seen in the proband of TD-1 (III-01) who had evidence for coronary artery disease at 38 years. This is particular noteworthy as TD patients, in addition to exhibiting significantly
5 reduced HDL, also have low LDL cholesterol, and yet they develop atherosclerosis despite this. This highlights the importance of HDL intracellular transport as an important mechanism in atherogenesis. There is significant evidence that heterozygotes for TD are also at increased risk for premature vascular disease (Schaefer *et al.*, Ann. Int. Med. 93:261-266, 1980; Serfaty-Lacrois *et al.*,
10 Atherosclerosis 107:85-98, 1994). There is also preliminary evidence for premature atherosclerosis in some probands with FHA (e.g., the proband in FHA-2 (III-01) had a coronary artery bypass graft at 46 years while the proband in FHA-3 had evidence for CAD around 50 years of age. The TD-1 proband had more severe efflux deficiency than the TD-2 proband. Interestingly, the TD-2 proband
15 had no evidence for CAD by 62 when he died of unrelated causes, providing preliminary evidence for a relationship between the degree of cholesterol efflux (mediated in part by the nature of the mutation) and the likelihood of atherosclerosis.

The *ABCI* gene plays a crucial role in cholesterol transport and, in
20 particular, intracellular cholesterol trafficking in monocytes and fibroblasts. It also appears to play a significant role in other tissues such as the nervous system, GI tract, and the cornea. Completely defective intracellular cholesterol transport results in peripheral neuropathy, corneal opacities, and deposition of cholesterol esters in the rectal mucosa.

25 HDL deficiency is heterogeneous in nature. The delineation of the genetic basis of TD and FHA underlies the importance of this particular pathway in

intracellular cholesterol transport, and its role in the pathogenesis of atherosclerosis. Unraveling of the molecular basis for TD and FHA defines a key step in a poorly defined pathway of cholesterol efflux from cells and could lead to new approaches to treatment of patients with HDL deficiency in the general population.

HDL has been implicated in numerous other biological processes, including but not limited to: prevention of lipoprotein oxidation; absorption of endotoxins; protection against *Trypanosoma brucei* infection; modulation of endothelial cells; and prevention of platelet aggregation (see Genest *et al.*, J. Invest. Med. 47: 31-42, 1999, hereby incorporated by reference). Any compound that modulates HDL levels may be useful in modulating one or more of the foregoing processes. Our previous discovery that ABC1 functions to regulate HDL levels links, for the first time, ABC1 with the foregoing processes.

With the identification of the ABC1 protein as a key initiator of the efflux pathway, it has now been possible to directly examine the relationship between efflux, HDL, triglyceride level, and CAD. We have characterized the phenotypes of heterozygotes for several mutations in the ABC1 gene in a large cohort where diagnosis has been made by mutation identification. Furthermore, the phenotype of the heterozygotes was compared to that of unaffected family members, enabling the results to be controlled, at least in part, for other genetic and environmental influences. In contrast, prior studies in obligate heterozygotes have been limited to small numbers, often within a single family, and thus restricted in the ability to analyze the phenotypic expression with multiple mutations over a range of ages.

A cohort of 77 individuals heterozygous for multiple mutations in the ABC1 gene were identified, enabling the characterization of 13 ABC1 mutations in 11 families (5 TD, 6 FHA). The ABC1 heterozygotes have an approximate

50% decrease in HDL cholesterol and apoAI, and a mild but significant decrease in apoAII. In addition, ABC1 heterozygotes have increased triglycerides, but in contrast to TD patients, have no significant change in total or LDL cholesterol. The changes in HDL, apoAI, and triglycerides were gene-dose dependent, suggesting that they are directly related to ABC1 function. Furthermore, heterozygotes have an over three-fold increased risk of developing CAD, and younger average age-of-onset compared to unaffected individuals. Further, the heterozygotes with the most severe deficiency in efflux had a higher frequency and greater severity of CAD. Interestingly, the severity of the phenotype observed in the heterozygotes appeared to be mutation-dependent, but there was no obvious relationship between the site of mutation and the phenotype. There was a trend toward lower HDL in carriers of severe mutations that caused truncations or null alleles than in carriers of missense mutations. One notable exception is the M1091T missense mutation which had the most severe phenotype, with marked reductions in HDL cholesterol and efflux in affected family members, suggesting that this mutation may act in a dominant-negative fashion, down-regulating the function of the wild-type allele. Another interesting finding is the small cluster of mutations at the very C-terminal region of the protein, which suggests that this region is critical for ABC1 function.

The severe HDL deficiency in ABC1 heterozygotes suggests that residual cholesterol efflux is the major determinant of HDL cholesterol levels. Here we demonstrated a strong correlation between cholesterol efflux and HDL cholesterol levels. From the regression equation of mean HDL on efflux, each 8% increase in relative efflux is predicted to be associated with a 0.1 mmol/L increase in HDL cholesterol levels. For example, to effect a 30% increase in HDL cholesterol in a 40 year old male, it would require a 50% increase in ABC1 mediated cholesterol

efflux. Although these numbers may not directly extrapolate to what is observed in a general population where other genetic and environmental factors have not been controlled for, these data nonetheless suggest that relatively small changes in ABC1 function may have a significant impact on plasma HDL cholesterol levels.

Furthermore, the data presented here suggest that variations in efflux due to variations in ABC1 function directly reflect not only plasma HDL cholesterol levels but also triglyceride levels and CAD susceptibility, thus providing direct validation of the reverse cholesterol transport hypothesis and validation of ABC1 as a therapeutic target to raise HDL cholesterol, lower triglyceride levels, and protect against atherosclerosis.

The phenotype in ABC1 heterozygotes is also age-modulated. From 20 years of age in members of the control cohort, there is a small but definite increase in HDL with advancing age that is obviously absent in the heterozygotes. One explanation for this finding is that there is normally an age-related increase in ABC1 function, which is not seen in heterozygotes, perhaps because the remaining functioning allele has already been maximally up-regulated secondary to an increase in intracellular cholesterol. This lack of age-related increase in ABC1 function in heterozygotes would exaggerate the difference in HDL levels between heterozygotes and control individuals in older age groups. There is some evidence for an age-modulated increase in expression of ABC transporters (Gupta, Drugs Aging 7:19-29, 1995). Further, evidence of a potential age-related increase in ABC1 function comes from the observation that the percentage of apoAI found in the pre β_1 subfraction of HDL, the predominant cholesterol acceptors, decreases with age, suggesting increased formation of mature α -migrating HDL with age.

In a Regression Growth Evaluation Statin Study, carriers of the R219K ABC1 mutation were found to have significantly lower triglyceride levels that

individuals without this mutation. This result suggests that compounds that bind near Arg219 in wild-type ABC1 or otherwise mimic the function provided by Lys219 in the R219K ABC1 variant may lower triglyceride levels, and thus decrease risk of CAD. In addition, carriers of the V399A ABC1 variant had higher HDL levels and fewer coronary events than individuals without this variant. Thus, compounds that bind near Val399 in wild-type ABC1 or mimic the function provided by Ala399 in the V399A ABC1 variant may increase cholesterol levels and decrease risk of CAD. Determining the presence or absence of the R219K or V399A ABC1 variants in individuals may be useful in selecting therapies (such as HDL-lowering, triglyceride-raising, or anti-CAD therapies) for these subjects.

The following examples are to illustrate the invention. They are not meant to limit the invention in any way.

Analysis of TD Families

Studies of cholesterol efflux

Both probands had evidence of marked deficiency of cholesterol efflux similar to that previously demonstrated in TD patients. TD-1 is of Dutch descent while TD-2 is of British descent.

Linkage analysis and establishment of a physical map

Multiple DNA markers were genotyped in the region of 9q31 to which linkage to TD had been described (Rust *et al.*, Nat. Genet. 20, 96-98, 1998). Two point linkage analysis gave a maximal peak LOD score of 6.49 at D9S1832 with significant evidence of linkage to all markers in a ~10 cM interval. Recombination with the most proximal marker, D9S1690 was seen in II-09 in Family TD-1, providing a centromeric boundary for the disease gene. Multipoint

linkage analysis of these data did not increase the precision of the positioning of the disease trait locus.

A physical map spanning approximately 10 cM in this region was established with the development of a YAC contig. In addition, 22 other polymorphic multi-allelic markers which spanned this particular region were mapped to the contig, and a subset of these were used in construction of a haplotype for further analysis.

While the family of Dutch decent did not demonstrate any consanguinity, the proband in TD-2 was the offspring of a first-cousin consanguineous marriage. We postulated, therefore, that it was most likely that this proband would be homozygous for the mutation while the proband in the Dutch family was likely to be a compound heterozygote. The Dutch proband shows completely different mutation bearing haplotypes, supporting this hypothesis.

The TD-2 proband was homozygous for all markers tested distal to D9S127 but was heterozygous at D9S127 and DNA markers centromeric to it. This suggested that the gene for TD was likely located to the genomic region telomeric of D9S127 and encompassed by the markers demonstrating homozygosity.

Mutation detection

Based on the defect in intracellular cholesterol transport in patients with TD, we reviewed the EST database for genes in this region which might be relevant to playing a role in this process. The *ABCI* transporter gene had previously been mapped to 9q31, but its precise physical location had not been determined (Luciani *et al.*, Genomics 21:150-159, 1994). The *ABCI* gene is a member of the ATP binding cassette transporters which represents a super family of highly conserved proteins involved in membrane transport of diverse substrates

including amino acids, peptides, vitamins and steroid hormones (Luciani *et al.*,
 Genomics 21:150-159, 1994; Dean *et al.*, Curr. Opin. Gen. Dev. 5:779-785, 1995).
 Primers to the 3' UTR of this gene mapped to YACs spanning D9S306 (887-B2
 and 930-D3) compatible with it being a strong candidate for TD. We initiated
 5 large scale genomic sequencing of BACs spanning approximately 800 kb around
 marker D9S306 (Research Genetics RPCI-11 BACs: 4I8, 31J20, 47O19, and
 179G21, which are publically available from Research Genetics which is located at
 2130 Memorial Parkway, Huntsville, Alabama 35801) BACs 4I8, 31J20, 47O19,
 and 179G21 are identical to previously described BACs 269, 274, 279 and 291,
 10 respectively (U.S.S.N. 09/526,193; U.S.S.N. 60/124,702; U.S.S.N. 60/138,048;
 U.S.S.N. 60/139,600; U.S.S.N. 60/151,977; Brooks-Wilson *et al.*, Nat. Genet.
 22:336-345, 1999). The *ABCI* gene was revealed encompassing 49 exons and a
 minimum of 75 Kb of genomic sequence. In view of the potential function of a
 gene in this family as a cholesterol transporter, its expression in fibroblasts and
 15 localization to the minimal genomic segment underlying TD, we formally assessed
 ABC1 as a candidate.

Patient and control total fibroblast RNA was used in Northern blot analysis
 and RT-PCR and sequence analyses. RT-PCR and sequence analysis of TD-1
 revealed a heterozygous T to C substitution in the TD-1 proband, which would
 20 result in a substitution of arginine for cysteine at a conserved residue between
 mouse and man. This mutation, confirmed by sequencing exon 31 of the ABC1
 gene, exhibited complete segregation with the phenotype on one side of this
 family. This substitution creates a *HgaI* site, allowing for RFLP analysis of
 amplified genomic DNA and confirmation of the mutation. The point mutation in
 25 exon 31 was not seen on over 200 normal chromosomes from unaffected persons
 of Dutch decent, and 250 chromosomes of Western European decent, indicating it

is unlikely to be a polymorphism. Northern blot analysis of fibroblast RNA from this patient, using a cDNA encompassing exons 2 to 50 of the gene, revealed a normal sized ~8 Kb transcript and a truncated mutant transcript which was not visible in control RNA or in RNA from other patients with HDL deficiency.

5 Additionally, Northern blot analysis using clones encompassing discrete regions of the cDNA revealed that the mutant transcript was detected with a cDNA encompassing exons 2 to 50, 2 to 42, 2 to 23, much more faintly with a probe spanning exon 24 to 30, and not seen with probes encompassing exons 31 to 43 or a probe spanning exons 31 to 50. This was repeated on multiple filters with
10 control RNA, RNA from other patients with HDL deficiency and the other TD proband, and only in TD-1 was the truncated transcript observed. Sequence analysis of the coding region did not reveal an alteration in sequence that could account for this finding. Furthermore, DNA analysis by Southern blot did not reveal any major rearrangements. Completion of exon sequencing in genomic
15 DNA showed that this mutation was a G to C transversion at position (+1) of intron 24, affecting a splice donor site and causing aberrant splicing.

RT-PCR analysis of fibroblast RNA encoding the *ABCI* gene from the proband in TD-2 revealed a homozygous nucleotide change of A to G at nucleotide 1864 in exon 14, resulting in a substitution of arginine for glutamine at
20 residue 597, occurring just proximal to the first predicted transmembrane domain of ABC1 at a residue conserved in mouse and as well as a *C. elegans* homolog. This mutation creates a second *AciI* site within exon 14. Segregation analysis of the mutation in this family revealed complete concordance between the mutation and the low HDL phenotype as predicted. The proband in TD-2 is homozygous
25 for this mutation, consistent with our expectation of a disease causing mutation in this consanguineous family.

Analysis of FHA families

Linkage analysis and refinement of the minimal genomic region containing the gene for FHA

Data from microsatellite typing of individual family members from the four
5 pedigrees of French Canadian origin were analyzed. A maximum LOD score of
9.67 at a recombination fraction of 0.0 was detected at D9S277 on chromosome
9q31. Thereafter, 22 markers were typed in a region spanning 10 cM around this
locus in these families. The frequency for these markers were estimated from a
sample of unrelated and unaffected subjects of French ancestry.

10 TD and FHA have thus far been deemed distinct with separate clinical and
biochemical characteristics. Even though the genes for these disorders mapped to
the same region, it was uncertain whether FHA and TD were due to mutations in
the same gene or, alternatively, due to mutations in genes in a similar region.
Refinement of the region containing the gene for FHA was possible by examining
15 haplotype sharing and identification of critical recombination events. Seven
separate meiotic recombination events were seen in these families, clearly
indicating that the minimal genomic region containing the potential disease gene
was a region of approximately 4.4 cM genomic DNA spanned by marker D9S1690
and D9S1866. This region is consistent with the results of two point linkage
20 analysis which revealed maximal LOD scores with markers D9S277 and D9S306
and essentially excluded the region centromeric to D9S1690 or telomeric to
D9S1866. An 8th meiotic recombination event further refined the FHA region to
distal to D9S277.

As described herein, the *ABCI* gene mapped within this interval. The
25 overlapping genetic data strongly suggested that FHA may in fact be allelic to TD.
Utilization of sets of genetic data from FHA and TD provided a telomeric

boundary at D9S1866 (meiotic recombinant) and a centromeric marker at D9S127 based on the homozygosity data of TD-2. This refined the locus to approximately 1 mb between D9S127 and D9S1866. The *ABC1* gene mapped within this minimal region.

5

Mutation detection in FHA

Mutation assessment of the *ABC1* gene was undertaken in FHA-1. Using primers that spanned overlapping segments of the mRNA we performed RT-PCR analysis and subjected these fragments to mutational analysis. A deletion of three nucleotides is evident in the RT-PCR sequence of FHA-1 III.01, resulting in a loss of nucleotides 2151-2153 and deletion of a leucine (Δ L693) at amino acid position 693. This leucine is conserved in mouse and *C. elegans*. The alteration was detected in the RT-PCR products as well as in genomic sequence from exon 15 specific amplification. This mutation results in a loss of an *Eae*I restriction site. Analysis of genomic DNA from the family indicated that the mutation segregated completely with the phenotype of HDL deficiency. The loss of the *Eae*I site results in a larger fragment being remaining in persons heterozygous for this mutation. This mutation maps to the first putative transmembrane domain of ABC1 and was not seen in 130 chromosomes from persons of French Canadian descent nor seen in over 400 chromosomes from persons of other Western European ancestry.

A mutation has also been found in patient genomic DNA in pedigree FHA-3 from Quebec. The alteration, a 6 bp deletion of nucleotides 5752-5757 within exon 42, results in a deletion of amino acids 1893 (Glu) and 1894 (Asp). The deletion was detected as a double superimposed sequence starting from the point of the deletion, and was detected in sequence reads in both directions. The

deletion can be detected on 3% agarose or 10% polyacrylamide gels, and segregates with disease in FHA-3. It was not seen in 128 normal chromosomes of French-Canadian origin or in 434 other control chromosomes. Amino acids 1893 and 1894 are in a region of the ABC1 protein that is conserved between human, mouse, and *C. elegans*, implying that it is of functional importance.

An additional mutation has been found in patient genomic DNA in pedigree FHA-2 from Quebec. The alteration, a C to T transition at position 6504, converts an arginine at position 2144 to a STOP codon, causing truncation of the last 118 amino acids of the ABC1 protein. This alteration segregates with disease in family FHA-2.

Functional relationship between changes in ABC1 transcript levels and cholesterol efflux

Antisense approaches were undertaken to decrease the ABC1 transcript and assess the effect of alteration of the transcript on intracellular cholesterol transport. The use of antisense primers to the 5' end of ABC1 clearly resulted in a decrease to approximately 50% of normal RNA levels. This would be expected to mimic in part the loss of function due to mutations on one allele, similar to that seen in heterozygotes for TD and patients with FHA. Importantly, reduction in the mRNA for the ABC1 gene resulted in a significant reduction in cellular cholesterol efflux, further establishing the role of this protein in reverse cholesterol transport and providing evidence that the mutations detected are likely to constitute loss of function mutations. Furthermore, these data support the functional importance of the first 60 amino acids of the protein. Antisense oligonucleotide AN-6 is directed to the novel start codon 5' to the one indicated in AJ012376.1; this antisense oligonucleotide effectively suppresses efflux.

Polymorphisms in ABC1 5' regulatory region and 5' UTR

Several polymorphisms in the 5' regulatory region of human ABC1 (SEQ ID NO: 1) have been identified (Fig. 4). Because of their location, it is likely that ABC1 gene expression will differ among humans having different promoter polymorphisms, and these individuals may also respond differently to the same drug treatment. Thus, using these newly-identified polymorphisms, one can tailor drug treatment depending on which polymorphism(s) is/are present in a patient. The presence or absence of particular ABC1 polymorphisms may also be used in determining an individual's predisposition to developing CVD.

The methods of the invention may be performed using the following materials and methods.

Biochemical studies

Blood is withdrawn in EDTA-containing tubes for plasma lipid, lipoprotein cholesterol, ApoAI, and triglyceride analyses, as well as storage at -80°C. Leukocytes are isolated from the buffy coat for DNA extraction.

Lipoprotein measurement is performed on fresh plasma as described elsewhere (Rogler *et al.*, Arterioscler. Thromb. Vasc. Biol. 15:683-690, 1995). Lipids, cholesterol and triglyceride levels are determined in total plasma and plasma at density $d < 1.006$ g/mL (obtained after preparative ultracentrifugation) before and after precipitation with dextran manganese. Apolipoprotein measurement is performed by nephelometry for ApoB and ApoAI.

Genomic clone assembly and physical map construction of the 9q31 region

Using the Whitehead Institute/MIT Center for Genome Research map as a reference, the genetic markers of interest at 9q31 were identified within YAC

contigs. Additional markers that mapped to the approximate 9q31 interval from public databases and the literature were then assayed against the YAC clones by PCR and hybridization analysis. The order of markers was based on their presence or absence in the anchored YAC contigs and later in the BAC contig. Based on the haplotype analysis, the region between D9S277 and D9S306 was targeted for higher resolution physical mapping studies using bacterial artificial chromosomes (BACs). BACs within the region of interest were isolated by hybridization of DNA marker probes and whole YACs to high-density filters containing clones from the RPCI-11 human BAC library.

Sequence retrieval and alignment

The human *ABC1* mRNA sequence was retrieved from GenBank using the Entrez nucleotide query (Baxeavanis *et al.*, A Practical Guide to the Analysis of Genes and Proteins, eds. Baxeavanis, A.D. & Ouellette, B.F.F. 98:120, 1998) as GenBank accession number AJ012376.1. The version of the protein sequence we used as wild-type (normal) was CAA10005.1.

We identified an additional 60 amino acids in-frame with the previously-believed start methionine. Bioinformatic analysis of the additional amino acids indicates the presence of a short stretch of basic amino acid residues, followed by a hydrophobic stretch, then several polar residues. This may represent a leader sequence, or another transmembrane or membrane-associated region of the ABC1 protein. In order to differentiate among the foregoing possibilities, antibodies directed to the region of amino acids 1-60 are raised against and used to determine the physical relationship of amino acids 1-60 in relation to the cell membrane. Other standard methods can also be employed, including, for example, expression of fusion proteins and cell fractionation.

The mouse *ABC1* sequence used has accession number X75926. It is very likely that this mouse sequence is incomplete, as it lacks the additional 60 amino acids described herein for human ABC1.

Version 1.7 of ClustalW was used for multiple sequence alignments with BOXSHADE for graphical enhancement (http://www.isrec.isb-sib.ch:8080/software/BOX_form.html) with the default parameter. A *Caenorhabditis elegans* ABC1 orthologue was identified with BLAST (version 2.08) using CAA1005.1 (see above) as a query, with the default parameter except for doing an organism filter for *C. elegans*. The selected protein sequence has accession version number AAC69223.1 with a score of 375, and an E value of 103.

Genomic DNA sequencing

BAC DNA was extracted from bacterial cultures using NucleoBond Plasmid Maxi Kits (Clontech, Palo Alto, CA). For DNA sequencing, a sublibrary was first constructed from each of the BAC DNAs (Rowen *et al.*, Automated DNA Sequencing and Analysis, eds. Adams, M.D., Fields, C. & Venter, J.C., 1994). In brief, the BAC DNA was isolated and randomly sheared by nebulization. The sheared DNA was then size fractionated by agarose gel electrophoresis and fragments above 2 kb were collected, treated with Mung Bean nuclease followed by T4 DNA polymerase and klenow enzyme to ensure blunt-ends, and cloned into *Sma*I-cut M13mp19. Random clones were sequenced with an ABI373 or 377 sequencer and fluorescently labeled primers (Applied BioSystems, Foster City, CA). DNASTar software was used for gel trace analysis and contig assembly. All DNA sequences were examined against available public databases primarily using BLASTn with RepeatMasker (University of Washington). The sequence of each of the assembled contigs is shown in Figs. 1A-D.

Reverse transcription (RT)-PCR amplification and sequence analysis

Total RNA was isolated from the cultured fibroblasts of TD and FHA patients, and reverse transcribed with a CDS primer containing oligo d(T)18 using 250 units of SuperScript II reverse transcriptase (Life Technologies, Inc.,
5 Rockville, MD) as described (Zhang *et al.*, J. Biol. Chem. 27:1776-1783, 1996). cDNA was amplified with Taq DNA polymerase using primers derived from the published human *ABC1* cDNA sequence (Luciani *et al.*, Genomics 21:150-159, 1994). Six sets of primer pairs were designed to amplify each cDNA sample, generating six DNA fragments which are sequentially overlapped covering 135 to
10 7014 bp of the full-length human *ABC1* cDNA. The nucleotides are numbered according to the order of the published human cDNA sequence (AJ012376.1). Primer pairs (1): 135-158 (f) and 1183-1199 (r); (2): 1080-1107 (f) and 2247-2273 (r); (3): 2171-2197 (f) and 3376-3404 (r); (4): 3323-3353 (f) and 4587-4617 (r); (5) 4515-4539 (f) and 5782-5811 (r); (6): 5742-5769 (f) and 6985-7014 (r). RT-
15 PCR products were purified by Qiagen spin columns. Sequencing was carried out in a Model 373A Automated DNA sequencer (Applied Biosystems) using Taq di-deoxy terminator cycle sequencing and Big Dye Kits according to the manufacturer's protocol.

20 Northern blot analysis

Northern transfer and hybridizations were performed essentially as described (Zhang *et al.*, J. Biol. Chem. 27:1776-1783, 1996). Briefly, 20 µg of total fibroblast RNA samples were resolved by electrophoresis in a denaturing agarose (1.2%; w/v) gel in the presence of 7% formaldehyde, and transferred to
25 nylon membranes. The filters were probed with ³²P-labeled human *ABC1* cDNA as indicated. Pre-hybridization and hybridizations were carried out in an

ExpressHyb solution (ClonTech) at 68°C according to the manufacturer's protocol.

Cell culture

5 Skin fibroblast cultures are established from 3.0 mm punch biopsies of the forearm of FHD patients and healthy control subjects as described (Marcil *et al.*, Arterioscler. Thromb. Vasc. Biol. 19:159-169, 1999).

Cellular cholesterol labeling and loading

10 The protocol for cellular cholesterol efflux experiments has been described in detail elsewhere (Marcil *et al.*, Arterioscler. Thromb. Vasc. Biol. 19:159-169, 1999). The cells are ³H-cholesterol labeled during growth and free cholesterol loaded in growth arrest.

Cholesterol efflux studies

15 Efflux studies are carried out from 0 to 24 hours in the presence of purified ApoAI (10 µg protein/mL medium). Efflux is determined as a percent of free cholesterol in the medium after the cells were incubated for specified periods of time. All experiments are preferably performed in triplicate, in the presence of
20 cells from one control subject and the cells from the study subjects to be examined.

Determination of genomic structure of the ABC1 gene

Most splice junction sequences were determined from genomic sequence generated from BAC clones spanning the ABC1 gene. More than 160 kb of
25 genomic sequence were generated. Genomic sequences were aligned with cDNA sequences to identify intron/exon boundaries. In some cases, long distance PCR

between adjacent exons was used to amplify intron/exon boundary sequences using amplification primers designed according to the cDNA sequence. The genomic sequence of human ABC1 is shown in Figs. 1A-D.

Analysis of ABC1 Heterozygotes

Identification of subjects

Subjects heterozygous for mutations in the ABC1 gene were individuals identified from the seven TD and FHA families previously described (Brooks-Wilson *et al.*, *supra*; Marcil *et al.*, *supra*). In addition, heterozygous individuals from three new Tangier disease families (TD3-5) and one new FHA kindred (FHA6) were included. The second mutation has not been identified in one of the TD kindreds (TD4); however, a marker immediately adjacent to ABC1 cosegregates with the low HDL phenotype. Individuals bearing the affected haplotype were considered heterozygotes. The presence or absence of mutations identified by genomic sequencing of probands from each family was subsequently confirmed by restriction fragment length polymorphism (RFLP) assays, to define heterozygous and unaffected individuals, respectively.

The control cohort consisted of unaffected members of the 11 families. These individuals share a genetic background with the heterozygotes, and environmental factors are expected to be similar amongst family members. Thus, many additional factors that influence HDL are controlled for, and the phenotypic differences between heterozygotes and unaffected individuals can be largely attributed to variation in ABC1 gene activity.

All subjects gave informed consent to their participation in this study, and the genetic analysis protocol was approved by the Ethics committees of the University of British Columbia, the Academic Medical Centre in Amsterdam and

the Clinical Research Institute of Montreal (IRCM).

Lipid and cholesterol efflux measurements

Lipid levels in ABCA1 heterozygotes were measured as previously
described (Brooks-Wilson *et al.*, *supra*; Marcil *et al.*, *supra*), at standardized lipid
clinics in Vancouver, Montreal and Amsterdam. LDL was calculated by the
method of Friedewald *et al.* (Clin. Chem. 18:499-502, 1972), modified to account
for lipid measurements in mmol/L.

Cellular cholesterol efflux from fibroblast cultures was measured as
previously described (Brooks-Wilson *et al.*, *supra*; Marcil *et al.*, *supra*). Each
experiment was performed in triplicate wells and averaged. Measurements are
reported as the percentage efflux in each subject relative to an average of at least
two healthy controls included within the same experiment. Individual experiments
were repeated at least twice, and the average relative efflux over all experiments
was used.

Statistics

In analysis of the heterozygotes, differences in mean baseline demographics
and lipid levels between groups were compared by Student's t-test. Comparisons of
frequency either between the male to female ratio or of distributions across various
percentile ranges were made using the chi-square test. Analyses of potential
interactions between affected status and either sex or BMI were performed using a
general linear model. Statistical analysis was performed using Prism (version
3.00, Graphpad Software) or Systat (version 8.0, SPSS Inc.). All values are
reported as mean \pm standard deviation.

Decreased HDL cholesterol and an increased risk for CAD in ABC1 heterozygotes

The analyzed cohort comprised 77 individuals from 11 families identified as heterozygous for mutations in the ABC1 gene. A comparison of mean lipid levels in heterozygotes with mean levels in all available unaffected family members (n=156) is presented in Fig. 6. Heterozygotes have an approximately 40-45% decrease in HDL and apoA-I and a mild (approximately 10%) decrease in apoA-II compared to unaffected family members. Mean triglycerides (TG) were increased by approximately 40% in heterozygotes compared to unaffected family members, and were further increased in TD patients. Unlike TD patients, there is no significant decrease in either total cholesterol (TC) or LDL cholesterol in heterozygotes, and apoB levels were not different in heterozygotes from controls. Mean HDL levels in carriers of each of the mutations were similarly reduced by approximately 40-50% compared to unaffected family members (Fig. 9).

The heterozygote phenotype was further examined by calculating the percentage of individuals falling within a given range of age and sex specific percentiles (based on LRC criteria (Heiss *et al.*, *supra*; Heiss *et al.*, Circulation 61:302-315, 1980). Much variability in the heterozygote phenotype was evident. As illustrated in Fig. 5A, although a significantly higher percentage of heterozygotes had HDL cholesterol less than the 5th percentile for age and sex compared to unaffected controls (65% vs. 5%, $p<0.0001$), 5% of the heterozygotes had HDL greater than the 20th percentile, with HDL ranging up to the 31st percentile for age and sex. Thus, in some individuals clearly the low HDL phenotype is less severe. A broad distribution of triglyceride (TG) levels was also evident (Fig. 5B). A significantly lower percentage of heterozygous individuals had TG below the 20th percentile for age and sex ($p=0.03$), and a significantly larger percentage had TG >80th percentile ($p=0.005$) compared to unaffected

family members, but substantial overlap between the two distributions was seen.

Another important question is whether individuals heterozygous for ABC1 mutations are at an increased risk of developing coronary artery disease (CAD). In our large cohort, symptomatic vascular disease was over three times as frequent in the adult heterozygotes as in unaffected family members (Fig. 6). The forms of vascular disease were generally more severe in the heterozygotes than in their unaffected family members (Fig. 7). Heterozygotes had myocardial infarctions (five, one fatal) and severe vascular disease requiring multiple interventions, whereas in unaffected individuals, CAD was manifest as angina in two cases and as a transient ischemic attack at the age of 80 in another. Furthermore, the mean age-of-onset was on average a decade younger in heterozygotes compared to unaffected controls (Fig. 6)

Cholesterol efflux, HDL level, and CAD in ABC1 heterozygotes

We have previously shown that individuals heterozygous for ABC1 mutations have decreased cholesterol efflux. In the present study, the relationship between cholesterol efflux levels, HDL cholesterol levels, and CAD was further assessed. Relative cholesterol efflux in individuals heterozygous for an ABC1 mutation was plotted against the mean HDL cholesterol levels observed in the carriers of that mutation, expressed as a percentage of the unaffected members within that family (Fig. 8). Cholesterol efflux levels associated with each mutation strongly predict the corresponding HDL cholesterol levels in the families, accounting for 82% of the variation in HDL cholesterol ($r^2=0.82$, $p=0.005$). Furthermore, in one large family (FHA2), where efflux has been measured in three independent heterozygotes, an r^2 value of 0.81 was obtained when individual plasma HDL levels were plotted against individual efflux measurements. Using

the regression equation of mean HDL levels in the heterozygotes on the efflux level of the heterozygous carrier ($p=0.02$), the relationship between expected changes in ABC1 efflux activity and HDL levels was estimated. Based on this analysis, we predict that each 8% change in efflux levels would be associated with a 0.1 mmol/L change in HDL cholesterol.

Relative cholesterol efflux levels are also related to CAD within the family. Families with clearest evidence for premature CAD had individuals with the lowest cholesterol efflux (bold on Fig. 8 and Fig. 9). These data suggest that the level of residual ABC1 function is a critical determinant of both HDL cholesterol levels and susceptibility to CAD.

Comparison of mutation type and location to the severity of phenotype in individuals heterozygous for ABC1 mutations

We have previously noted that the phenotypic presentation of our FHA heterozygotes was more severe than that of our TD heterozygotes. Furthermore, we initially noted more deletions and premature truncations of the protein in our FHA families than our TD families. Thus, as residual ABC1 activity is an important predictor of severity of the phenotype, the influence of the nature of the mutation on the phenotypic expression of mutations in the ABC1 gene was examined. Severe mutations which would be expected to result in a non-functional allele were defined as deletions or mutations that caused premature truncation of the protein (frameshifts and nonsense mutations) or that disrupted natural splicing of the protein. Missense mutations, on the other hand, result in the change of only a single amino acid and may result in a protein product that still retains partial activity.

Lipid levels were compared in heterozygous carriers of severe and missense

mutations. While there was a trend to decreased HDL levels in carriers of severe compared to missense mutations, this trend did not reach significance (Fig. 10). A range of HDL levels in individual missense and severe mutations were observed (Fig. 9). The M1091T missense mutation is the most severe mutation in terms of effects on efflux and HDL levels, with a more severe phenotype than even early truncations of the protein (e.g. R909X).

The site of mutation (e.g. N-terminal or C-terminal) within the ABC1 protein did not influence the phenotype (Fig. 11). The presence of CAD is seen in carriers of mutations in several domains of the protein. Patients with mutations on both alleles manifest with splenomegaly alone or in association with CAD (TD1). Thus the phenotype appears to be mutation specific, and most likely dependent on remaining ABC1 function of the wild-type allele and residual function of the mutant allele, similar to what has been shown for mutations in ABCR, a close homologue of ABC1 (van Driel *et al.*, Ophthalmic Genet. 19:117-122, 1998).

Relationship between phenotype of mutations in the ABC1 gene and age

One factor influencing phenotypic expression that became apparent in the families was age. We first characterized the effect of age on the HDL levels of individuals in two previously reported families (Marcil *et al.*, *supra*) (Figs. 12A and 12B). In family FHA3, while heterozygous individuals in generations II and III all had HDL cholesterol levels less than the 5th percentile for age and sex, those in generation IV had a much more variable phenotype, with HDL cholesterol ranging up to the 20th percentile. In family FHA1, the same pattern was observed.

The distribution of individuals less than 30 years old across HDL percentile ranges was compared to the corresponding distribution of individuals 30-70 years old across HDL percentile ranges (Fig. 13). A significantly larger percentage of

individuals 30-70 years old had HDL cholesterol less than the 5th percentile compared to the percentage of individuals less than 30 years old with HDL cholesterol less than the 5th percentile. Mean HDL decreases in heterozygotes greater than 30 years of age compared to those less than 30 years of age; in contrast, there is no significant change in unaffected controls (Fig. 14). Similar results are seen in males and females separately and in both pre- and post-menopausal ages in women (Figs. 15A and 15B). Triglyceride levels increase with age in both heterozygotes and unaffected family members.

Assessment of the influences of gender and BMI on the phenotypic expression of ABC1 mutations

Females are known to have elevated HDL and decreased triglyceride levels compared to males. Thus, the phenotype of ABC1 heterozygotes was analyzed to determine whether the phenotype was influenced by gender. HDL cholesterol is significantly lower than unaffected controls in both heterozygous males and females (0.70 ± 0.24 versus 1.21 ± 0.29 , $p < 0.0001$; 0.76 ± 0.25 versus 1.41 ± 0.38 , $p < 0.0001$, respectively). This was reflected in decreased apoAI (0.92 ± 0.27 versus 1.36 ± 0.22 , $p < 0.0001$; 0.92 ± 0.36 versus 1.49 ± 0.28 , $p < 0.0001$ in males and females, respectively), and a trend towards a mild decrease in apoAII in both males and females compared to unaffected family members (0.35 ± 0.08 versus 0.40 ± 0.09 , $p = 0.08$; 0.35 ± 0.09 versus 0.39 ± 0.07 , $p = 0.06$, respectively). Triglycerides are higher in both male (2.07 ± 2.16 versus 1.30 ± 1.30 , $p = 0.02$) and female (1.34 ± 0.86 versus 1.09 ± 0.63 , $p = 0.08$) heterozygotes compared to unaffected family members. The difference in HDL between males and females was reduced in heterozygotes compared to controls ($p = 0.11$), while the difference in triglycerides was increased compared to controls ($p = 0.13$).

Another factor known to influence HDL and triglyceride levels is BMI. The entire cohort was divided into tertiles of BMI, and the mean HDL and triglyceride levels of heterozygotes and unaffected individuals by BMI tertile are shown in Figs. 16A and 16B. BMI had a significant effect on both HDL and triglycerides in both heterozygotes and controls ($p < 0.0001$). The effect of BMI on HDL-C and triglyceride levels was more severe in heterozygotes for ABC1 than in controls, being evident at lower BMIs (mid-tertile) in heterozygotes. A raised BMI was more obviously associated with changes in HDL and triglyceride levels in heterozygotes compared to controls. However, neither effect reached statistical significance. HDL was reduced in heterozygotes compared to controls in all BMI tertiles ($p < 0.0001$ in each tertile). While triglyceride levels were increased in all BMI tertiles in heterozygotes compared to unaffected family members, this difference was only significant in the middle BMI tertile ($p = 0.009$).

Analysis of ABC1 SNPs from REGRESS Study

Identification of SNPs

SNPs in the ABC1 gene were identified during the complete genomic sequencing of 14 unrelated probands with low HDL-C (Brooks-Wilson *et al.*, *supra* 1999; Marcil *et al.*, *supra* 1999). Variants that were identified within the low HDL families that did not co-segregate with the low HDL phenotype or that were observed in unaffected individuals were assumed to be SNPs. Based on the sequencing of BAC clones spanning the entire ABC1 region (described above), sites identified as heterozygous or different from that found in sequenced individuals were also identified as polymorphisms. Sequence data was available from at least one control individual at all variant coding sites. The SNPs are numbered from the nucleotide described as position 1 (Pullinger *et al.*,

Biochemical and Biophysical Research Communications 271:451-455, 2000), naming the first exon number 1. As a standardized nomenclature for all variants, the "wild-type" allele (more frequent in the REGRESS population) is designated A, while the variant (less frequent) allele is designated B.

5

Subjects

To assess the effects of these SNPs on lipid levels and CAD, we studied a cohort of 804 men with proven coronary artery disease who participated in the Regression Growth Evaluation Statin Study (REGRESS), which has previously
10 been described in detail (Jukema *et al.*, Circulation 91:2528-2540, 1995). Briefly, study participants were required to have at least one coronary artery with a stenosis of more than 50% as assessed by coronary angiography, a plasma total cholesterol concentration of 4 to 8 mmol/L (155 to 310 mg/dL), and a plasma triglyceride concentration below 4 mmol/L (350 mg/dL). Phenotypic effects of the cSNPs
15 were examined in relationship to baseline lipid parameters.

Patients were randomly assigned to treatment with pravastatin (Pravachol, Bristol-Myers Squibb, Princeton, N.J.) or placebo for a period of two years. Computer-assisted quantitative coronary angiography was carried out at the start and at the end of the study as previously described (Jukema *et al.*, *supra* (1995)).
20 The baseline values and changes in the average mean segment diameter (MSD), which is a measurement of the average unobstructed diameter along the vessel, and in the minimal obstructive diameter (MOD), which is a measurement of the smallest unobstructed segment, were used as the primary measures of CAD. The MSD reflects diffuse changes of atherosclerosis, and the MOD reflects focal
25 atherosclerotic changes. Larger MSD and MOD measurements reflect less occlusion of the vessel, and a decrease in these parameters reflects progression of

coronary atherosclerosis. In addition, the prevalence of coronary events; defined as death, myocardial infarction, unscheduled coronary angioplasty or bypass surgery (PTCA, CABG), or stroke/transient ischemic attack; was examined.

Blood was collected from each patient at baseline, and DNA was extracted according to standard procedures. Several subsequent genetic studies have been performed on this cohort (Reymer *et al.*, Nature Genetics 10:28-34, 1995; Jukema *et al.*, Circulation 94:1913-1918, 1996; Kluijtmans *et al.*, Circulation 96:2573-2577, 1997; Kastelein *et al.*, Clinical Genetics 53:27-33, 1998; Kuivenhoven *et al.*, New England Journal of Medicine 338:86-93, 1998). The REGRESS and its DNA substudies were approved by all seven institutional review boards of the participating centers and by their medical ethics committees.

Additional Dutch subjects with low HDL and premature coronary artery disease were obtained from previously described populations (Kuivenhoven *et al.*, Arteriosclerosis, Thrombosis and Vascular Biology 17:560-568, 1997; Verhoff *et al.*, Atherosclerosis 141:161-166 1998; Franco *et al.*, British Journal of Hematology 102:1172-1175, 1998; Wittekoek *et al.*, Artherosclerosis 146:271-279, 1999; Franko *et al.*, British Journal of Hematology 104:50-54, 1999). Dutch control subjects were taken from a large population based study designed to assess the effects of various risk factors on CAD (Seidell *et al.*, International Journals of Obesity 19:924-927 1995; Kuivenhoven *et al.* Arteriosclerosis, Thrombosis and Vascular Biology 17:595-599, 1997). French Canadian subjects were a random sample of individuals. The South African Black and Cantonese cohorts have previously been described (Ehrenborg *et al.*, Arteriosclerosis, Thrombosis and Vascular Biology 17:2672-2678, 1997). All subjects gave informed consent.

cSNP screening

For each variant, a restriction enzyme whose cleavage pattern was altered by the variant was identified for development of an RFLP assay. If no suitable enzyme was found, a mismatch strategy was employed, whereby a single nucleotide mismatch was incorporated into the PCR primer, creating a restriction site in combination with either the wild-type or variant allele. The specific conditions of all assays are described in Fig. 17. All PCR reactions were carried out in 50 μ L volumes, in the presence of 1x PCR buffer and 1.5 μ M $MgCl_2$ (Life Technologies). Thermocycling parameters for all assays were as follows: 95°C for 3 min; 35 cycles of denaturation at 95°C for 10 seconds, annealing for 30 seconds at the temperature specified in Fig. 17, and elongation for 30 seconds at 72°C; and a final elongation at 72°C for 10 min. All digestions (15-20 μ L PCR product) were carried out in manufacturer's buffer (New England Biolabs) for 2 hours at the temperature specified by the manufacturer. As an example, the digest results for the R219K are shown in Fig. 18. A 177 base pair fragment with the A allele is not cut by *Eco*NI, whereas the B allele is digested to produce fragments of 107 and 70 base pairs. Heterozygous individuals thus display all three bands (177, 107, and 70 base pairs).

Genotyping with the TaqMan(r) assay

To facilitate the mass screening of some variants, TaqMan based polymerase chain reaction (PCR) assays (Holland *et al.*, Proc.Natl.Acad.Sci. 88 (16):7276-7280, 1991; Lee *et al.*, Nucleic Acids Research 21 (16):3761-3766, 1993) were developed for the detection of polymorphisms in the ABC1 gene. In this one-tube assay, two fluorogenic hybridization probes (one for each allele) are labeled with different fluorescent reporter dyes (FAM or TET) at their 5' terminus and a common quencher dye (TAMRA)

at their 3' terminus. These probes are cleaved by the 5' nuclease activity of Taq enzyme during PCR amplification. This cleavage separates the reporter from the quencher dye and generates an increase in reporter fluorescence. By using two different reporter dyes, cleavage of allele-specific probes can be detected in a single PCR. The difference in the measured fluorescence intensity between the two TaqMan probes allows for accurate allele calling.

PCR amplifications with flanking sets of primers (300 nM) in the presence of two TaqMan probes (25 nM each) and 4.5 mM MgCl₂ were performed using the following thermocycling protocol: initial denaturation at 96°C for 10 min, followed by 39 cycles of 96°C for 30 sec, 63°C for 1 min and 72°C for 15 sec, followed by a final extension at 72°C for 10 min. Each plate included controls (no DNA template) as well as standards of each known genotype. Fluorescence quantification and genotype determination were performed on a Perkin Elmer LS50B or ABI Prism 7700 Sequence Detector. The fluorescence from each reaction was normalized to the signal from the no-template controls.

Cellular cholesterol efflux studies

Cellular cholesterol efflux from fibroblast cultures was measured as described above in the “Analysis of ABC1 Heterozygotes” section.

Statistics

Within the REGRESS population, the baseline characteristics of the patients in the three genotypes (AA, AB, BB) was compared using one way analysis of variance and the chi-square test, where appropriate. In cases where the BB genotype was rare, the AA group was compared to the combined group AB+BB.

The cumulative coronary event incidence in these genotypic groups was compared using the logrank test. The relation between age and HDL level was investigated using a linear regression model, and the difference between genotypes with respect to the slopes of this regression line was tested using covariance analysis. In addition to this regression analysis, the R219K genotypes were compared among age-defined subgroups. Randomization to placebo and pravastatin was assessed by chi square analysis and was equivalent in all genotypic groups for all variants except the R1587K variant. For this variant, a lower proportion of carriers was randomized to pravastatin treatment. In addition, change in MOD, MSD, and prevalence of coronary events (the three variables measured during the trial and thus following randomization) were analyzed for the placebo and pravastatin subgroups separately. Similar genotypic effects for each of the variants were observed in the treatment subgroups (i.e. the pravastatin and placebo control groups).

All lipid levels are reported in mmol/L. All values are reported as mean \pm standard deviation.

Association of R219K polymorphism with reduced triglyceride levels and a decreased risk of CAD

The common R219K polymorphism results in the substitution of a lysine for an arginine at amino acid 219 of the ABC1 protein. The allele frequency of the variant, or "B", allele was 0.25, and its carrier frequency was 46.3%, as shown in Fig. 19.

The relationship between this polymorphism and CAD was examined. The B allele of the R219K polymorphism was associated with decreased baseline CAD. Both the MSD and the MOD, were significantly increased in an allele-dose

dependent fashion from AA to AB to BB (Fig. 20).

The angiographic data was paralleled by differences in clinical coronary events rates. A smaller percentage of individuals homozygous for the rarer B allele (BB) had had a myocardial infarction (MI) prior to the initiation of the trial. Carriers of the B allele exhibited a strong trend towards an increased prevalence of no events over the course of the study (Fig. 21, $p=0.07$).

The association of the R219K polymorphism with reduced CAD was further supported by the reduced frequency of this allele in the REGRESS cohort which was selected for CAD. In fact, the genotype frequencies observed for this variant are not consistent with Hardy-Weinberg equilibrium ($p=0.004$). There are fewer BB individuals and more AB individuals than would be expected (424 AA, 330 AB, and 36 BB individuals were observed compared to the expected 439 AA, 300 AB, and 51 BB individuals). As the REGRESS cohort was selected for men with CAD, the lack of Hardy-Weinberg equilibrium suggests that there was a preferential selection against BB individuals, consistent with the reduced CAD observed in this group.

There were no obvious differences in mean HDL-C levels between the genotypes in the groups as a whole (Fig. 22); however, triglycerides were significantly lower in the carriers of the B allele. This suggests that ABC1 function may also directly influence plasma triglyceride levels, and that this variant may be associated with a gain of ABC1 function. These findings are also consistent with the decreased CAD observed in carriers, suggesting that ABC1 modulation of triglyceride levels may be another mechanism whereby ABC1 activity influences risk of CAD.

To further explore the apparent lack of effect of this variant on HDL-C, the relationship between R219K genotype and HDL at various ages was examined. In

younger individuals, the carriers of the B allele had increased HDL-C compared to non-carriers (Fig. 23). Furthermore, while HDL-C increased with age in the AA individuals, heterozygous carriers showed a much milder increase, and HDL cholesterol decreased with age in homozygous carriers (Fig. 24). In the AA individuals, HDL cholesterol is positively correlated with age ($p < 0.001$). In contrast, this relation was not apparent in individuals heterozygous for this variant, and HDL cholesterol was negatively correlated with age in the BB homozygotes, although neither the correlation for AB nor the correlation for BB individuals was statistically different from zero (Fig. 25A). The age-related increase in HDL-C in AA individuals is not maintained in carriers of the B allele (p value comparing slopes=0.04). Thus, the decreased CAD in carriers of the R219K may also be related to the fact that for the majority of their lifetime, carriers have had increased HDL-C compared to non-carriers.

The changes in HDL-C with age in the various R219K genotypes are mirrored by similar trends in the changes in cholesterol efflux with age (Fig. 25B). We have genotyped this variant in all individuals in whom we have measured efflux and who do not possess an ABC1 mutation. In AA individuals ($n=30$), cholesterol efflux increases with age, whereas in AB and BB individuals ($n=24$), efflux decreases with age (p -value comparing slopes=0.15). Thus, the differential age-related changes in HDL seen in the different R219K genotypes are consistent with similar functional changes in ABC1 activity.

From Fig. 23 it can be seen that triglyceride levels generally decrease with age, a finding seen in all R219K genotypes. The percent decrease in triglyceride levels was nearly half in carriers (9.3%) compared to non-carriers (17.3%, $p=0.07$). Differences between the genotypes are also observed in the MSD and MOD. In the non-carriers, MSD and MOD measurements decrease significantly

with age, reflecting increased atherosclerosis in the older individuals (Figs. 23 and 26). In contrast, in carriers of the R219K variant, these measurements do not significantly change with age. Thus, vascular disease progresses much more slowly with age in carriers of the R219K variant compared to non-carriers.

5 Populations of Asian and African origins have been shown to have increased HDL-C, decreased triglyceride levels, and decreased risk of CAD compared to Caucasian populations (Tyroler *et al.*, Circulation 62 (Suppl. IV):IV-99-IV-107, 1980; Tao *et al.*, International Journal of Epidemiology 21 (5):893-903, 1992; Brown *et al.*, Arteriosclerosis and Thrombosis 13:1139-1158. 10 1993; Adedeji, Tropical and Geographical Medicine 46 (1):23-26, 1994; Simon *et al.*, American Journal of Public Health 85 (12):1698-1702, 1995; Morrison *et al.*, Metabolism 47 (5):514-521, 1998), a finding paralleling the phenotypic effects of this variant. Thus, as SNP frequencies can often differ within different ethnic groups, the frequency of this variant within these two population groups was 15 examined (Fig. 27). This variant is seen much more commonly in individuals of either Cantonese or South African Black descent, in which it is the predominant allele. This suggests that the increased frequency of this variant may in part account for the increased HDL, decreased triglyceride levels, and decreased CAD observed in these populations compared to Caucasian populations.

20 *Effect of other cSNPs on plasma lipid levels and risk of CAD*

No significant differences in lipid levels or CAD compared to respective non-carriers have been observed for carriers of the T774P (n=4), K776N (n=3) or E1172D variants (n=34). Carriers of the V825I (n=103 AB, 4 BB) had no obvious 25 differences in lipid levels or baseline MSD or MOD. Carriers of the V825I variant did, however, have a significantly increased number of events during the trial

(44% versus 33%, $p=0.001$).

No carriers of the S1731C variant were detected in the REGRESS population, but this variant was found in one of our FHA families (FHA2). In individuals heterozygous for ABC1 mutations, this variant was associated with significantly decreased HDL-C (0.16 ± 0.04 , $n=2$ versus 0.64 ± 0.14 , $n=10$; $p=0.0009$ in carriers versus non-carriers). In unaffected family members, however, while carriers of the S1731C had lower HDL-C compared to non-carriers (1.03 ± 0.22 versus 1.09 ± 0.23), this was not statistically significant. The control individual in whom this variant was also seen (Fig. 19) had low plasma HDL-C (0.72 mmol/L).

Similar to carriers of the R219K, carriers of the V771M variant ($n=37$), had no difference in HDL-C compared to non-carriers; however, a marginally significant interaction between age and genotype on HDL-C levels was noted ($p=0.05$). All but 2 carriers of the V771M variant are also carriers of the R219K variant. The interaction between age and genotype on HDL-C remains nearly significant when adjusted for R219K genotype ($p=0.11$), thus this variant may have an age effect independent of that which can be attributed to the R219K. A trend to less CAD (increased MOD) was observed in carriers of this variant compared to non-carriers (1.89 ± 0.38 versus 1.76 ± 0.35 , $p=0.13$).

For the I883M cSNP, homozygous BB individuals ($n=14$) have increased progression in MOD (mean change of 0.53 ± 0.79 versus 0.11 ± 0.25 , $p<0.001$). BB individuals had an events rate double that of the AA individuals (21.4% versus 10.6%) although this was not statistically significant ($p=0.19$). No difference was observed in mean lipid levels between the I883M genotypes. Furthermore, there were significantly more BB individuals than expected under Hardy-Weinberg equilibrium (14 BB, 86 AB, and 320 AA individuals observed compared to the expected 8 BB, 98 AB, and 314 AA individuals). As this cohort was selected for

individuals with CAD, this might suggest a preferential inclusion of those with the BB genotype. The association of this variant with CAD is further supported by the significantly increased frequency of this variant in the premature CAD population (odds ratio for CAD in carriers of this variant= 0.43, 95% confidence interval 0.22-0.85, $p=0.01$) (Fig. 19). These findings contrast with those of a very recent report which suggests that the homozygous carriers of this cSNP have increased HDL-C (Wang *et al.*, Arteriosclerosis, Thrombosis and Vascular Biology 20:1983-1989, 2000).

Carriers of the V399A (AB, $n=9$) had a trend to higher HDL-C (1.03 ± 0.28 versus 0.92 ± 0.23 , $p=0.15$) compared to individuals who were AA at this site ($n=540$). No events were observed in the AB group (compared to 14% in AA's, $p=NS$), and carriers had half the prevalence of a family history of CAD (22.2% versus 49.4%, $p=0.18$). Furthermore, consistent with this data, the carriers had a trend to increased baseline MOD (1.92 ± 0.32 versus 1.73 ± 0.35 , $p=0.13$) and to less progression in MSD (-0.05 ± 0.10 versus 0.08 ± 0.19 , $p=0.16$) during the trial. However, as the number of carriers of small, firm conclusions regarding the relationship of this variant to increased HDL-C and decreased CAD cannot be drawn.

Carriers of the R1587K variant (AB, BB) have decreased HDL-C compared to non-carriers in an allele-dose dependent trend (0.86 ± 0.16 , 0.91 ± 0.23 , and 0.94 ± 0.23 for BB, AB, and AA, respectively, $p=0.03$). No significant interaction with age was noted ($p=0.32$). Furthermore, on multiple regression including age, BMI, smoking, and triglyceride levels as covariates, the R1587K genotype is a significant predictor of HDL-C ($p=0.027$). However, no significant differences in CAD or events during the trial were evident in carriers compared to non-carriers.

Summary of Association Studies of ABC1 Gene Variants and HDL Levels or Cardiovascular Disease

The following polymorphisms have been examined for their effect on cholesterol regulation and the predisposition for the development of cardiovascular disease. The polymorphisms are numbered from the nucleotide described as position 1 (Pullinger *et al.* , *supra*), naming the first exon number 1.

Substitution of A for G at nucleotide 1051 (R219K). Carriers of this variant have reduced triglyceride levels, increased HDL cholesterol levels (particularly in younger individuals), and reduced CAD.

Substitution of C for T at nucleotide 1591 (V399A). This variant was associated with a trend towards increased HDL cholesterol in carriers.

Substitution of A for G at nucleotide 2706 (V771M). Carriers of this variant have been shown to have decreased CAD.

Substitution of C for A at nucleotide 2715 (T774P). This variant was seen less often in individuals with low HDL cholesterol levels or CAD than in controls.

Substitution of C for G at nucleotide 2723 (K776N). This variant has been found at a lower frequency (0.54% versus 1.89%) in a coronary artery disease population versus a control population of similar Dutch background.

Substitution of C for G at nucleotide 3911 (E1172D). This variant is seen at lower frequencies in individuals with low HDL and in some populations with premature coronary artery disease.

Substitution of A for G at nucleotide 5155 (R1587K). This variant is associated with decreased HDL cholesterol levels in carriers.

Substitution of G for C at nucleotide 5587 (S1731C). Two FHA individuals who have this variant on the other allele have much lower HDL cholesterol (0.155±0.025) than the FHA individuals in the family who do not have this variant

on the other allele (0.64 ± 0.14 , $p=0.0009$). This variant has also been found in one general population French Canadian control with HDL at the 8th percentile (0.92) and one French Canadian individual from a population selected for low HDL cholesterol levels and coronary disease (0.72).

5 *Substitution of G for A at nucleotide 2723 (I883M)*. This variant has been seen at a much higher frequency in individuals of Dutch ancestry with premature coronary artery disease. Furthermore, homozygous carriers of this variant have significantly increased CAD progression compared to non-carriers.

10 *Substitution of A for G at nucleotide 2868 (V825I)*. Carriers of this variant had significantly more CAD events than individuals who do not have this variant.

Substitution of C for G at nucleotide -191. Homozygous carriers of this variant have a three-fold increase in the frequency of coronary events (33.3% versus 11.2%, $p=0.003$) and a nearly double frequency of a positive family history of CAD (73.3% versus 47.7%, $p=0.01$).

15 *Substitution of G for C at nucleotide -17*. Carriers of this variant have significantly decreased coronary events (12.3% versus 18.2%, $p=0.04$) and a significantly decreased incidence of myocardial infarction (heart attack, 43.6% versus 52.8%, $p=0.02$).

20 *Substitution of T for C at nucleotide 69*. Carriers of this variant have increased CAD progression compared to non-carriers.

Substitution of G for C at nucleotide 127. Carriers of this variant have a trend towards decreased progression of CAD compared to non-carriers.

Insertion of CCCT at nucleotide -1163 in intron 1. Carriers of this variant have a trend to lower HDL cholesterol levels.

25 *Substitution of G for A at nucleotide -1095 in intron 1*. Homozygous carriers of this variant have a trend towards decreased HDL cholesterol and

increased triglyceride levels compared to non-carriers.

Substitution of A for G at nucleotide -1027 in intron 1. Carriers of this variant are also carriers of the G(-720)A. Thus the effects attributed to that variant may also be attributed to carriers of this variant.

5 *Substitution of A for G at nucleotide -720 in intron 1.* Homozygous carriers of this variant had a trend towards an increased frequency of a positive family history of myocardial infarction.

10 *Substitution of C for A at nucleotide -461 in intron 1.* Carriers of this variant are also carriers of the A(-362)G. Thus the effects attributed to that variant may also be attributed to carriers of this variant.

Substitution of G for A at nucleotide -362 in intron 1. Carriers of this variant have decreased triglyceride levels compared to non-carriers.

Insertion of G at nucleotide 319. Carriers of this variant have increased CAD compared to non-carriers.

15 *Substitution of G for C at nucleotide 378.* Carriers of this variant are also carriers of the InsG319. Thus the effects attributed to that variant may also be attributed to carriers of this variant.

Functional Role of LXRE Binding Sites in ABC1 Genomic Sequence

20 The functional role of three of the LXRE consensus binding sites identified in the ABC1 genomic sequence of SEQ ID NO: 1 was confirmed using standard gel shift assay experiments. Briefly, the sequences of the LXRE consensus binding sites located at -4389 and -1641 of the promoter region, +4 of exon 1, and -7670 and -7188 of 3' intron 1 (Figure 3) were tested in gel shift assays using the
25 cyp7 LXRE described by Lehmann *et al.* as a positive control (Lehmann *et al.*, J. Biol. Chem. 272:3137-3140, 1997).

In the first gel shift assay, binding of the labeled cyp7 LXRE probe was competed by a 400-fold excess of cold cyp7 probe, -4389 probe, -1641 probe, +4 probe, or -7670 probe. The signal disappeared completely in all cases.

For the next assay, the -4389, -1641, +4, -7670 and -7188 probes were labeled and binding to LXR-RXR complex was assayed. A signal similar to that of the positive control was observed for each of ABC1 LXRE probes, although the signal was slightly more intense for the +4 probe.

A competition assay was also performed using lesser quantities of each cold probe, i.e. 5-fold, 25-fold, and 50-fold more cold probe than labeled cyp7 probe.

There was a dose dependant decrease in the signal for each of the probes. This decrease was more significant for the +4 and -7670 probes. Moreover, the signal was not modified by competition with a cold Dr2 like probe, suggesting that the competition effect is indeed specific.

Thus, each of the tested potential LXRE binding sites seem to bind an *in vitro* LXR-RXR heterodimer. The LXRE binding site at +4 in exon 1 appears to have the highest affinity, closely followed by the LXRE binding site at -7670 in 3' intron 1.

Agonists and Antagonists

Useful therapeutic compounds include those which modulate the expression, activity, or stability of ABC1. To isolate such compounds, ABC1 expression, biological activity, or regulated catabolism is measured following the addition of candidate compounds to a culture medium of ABC1-expressing cells. Alternatively, the candidate compounds may be directly administered to animals (for example mice, pigs, or chickens) and used to screen for their effects on ABC1 expression.

In addition its role in the regulation of cholesterol, ABC1 also participates in other biological processes for which the development of ABC1 modulators would be useful. In one example, ABC1 transports interleukin-1 β (IL-1 β) across the cell membrane and out of cells. IL-1 β is a precursor of the inflammatory response and, as such, inhibitors or antagonists of ABC1 expression or biological activity may be useful in the treatment of any inflammatory disorders, including but not limited to rheumatoid arthritis, systemic lupus erythematosus (SLE), hypo- or hyper- thyroidism, inflammatory bowel disease, and diabetes mellitus. In another example, ABC1 expressed in macrophages has been shown to be engaged in the engulfment and clearance of dead cells. The ability of macrophages to ingest these apoptotic bodies is impaired after antibody-mediated blockade of ABC1. Accordingly, compounds that modulate ABC1 expression, stability, or biological activity would be useful for the treatment of these disorders.

ABC1 expression is measured, for example, by standard Northern blot analysis using an *ABC1* nucleic acid sequence (or fragment thereof) as a hybridization probe, or by Western blot using an anti-ABC1 antibody and standard techniques. The level of ABC1 expression in the presence of the candidate molecule is compared to the level measured for the same cells, in the same culture medium, or in a parallel set of test animals, but in the absence of the candidate molecule. ABC1 activity can also be measured using the cholesterol efflux assay.

Transcriptional Regulation of ABC1 Expression

ABC1 mRNA is increased approximately 8-fold upon cholesterol loading. This increase is likely controlled at the transcriptional level. Using the genomic sequence described herein, one can identify transcription factors that bind to the 5' regulatory sequence by performing, for example, gel shift assays, DNase

protection assays, or *in vitro* or *in vivo* reporter gene-based assays. The identified transcription factors are themselves drug targets. In the case of ABC1, drug compounds that act through modulation of transcription of ABC1 could be used for HDL modulation, triglyceride modulation, atherosclerosis prevention, and the treatment of cardiovascular disease. For example, using a compound to inhibit a transcription factor that represses ABC1 would be expected to result in up-regulation of ABC1 and, therefore, up-regulation of HDL cholesterol levels and down-regulation of triglyceride levels. In another example, a compound that increases transcription factor expression or activity would also increase ABC1 expression, increase HDL levels, and decrease triglyceride levels.

Transcription factors known to regulate other genes in the regulation of apolipoprotein genes or other cholesterol- or lipid-regulating genes are of particular relevance. Such factors include, but are not limited to, the steroid response element binding proteins (SREBP-1 and SREBP-2), and the PPAR (peroxisomal proliferation-activated receptor), RXR, and LXR transcription factors. Several consensus sites for certain elements are present in the sequenced region 5' to the ABC1 gene (Fig.3) and thus are likely to modulate ABC1 expression. For example, LXRs may alter transcription of ABC1 by mechanisms including heterodimerization with retinoid X receptors (RXRs) and then binding to specific response elements (LXREs). Examples of such LXRs include LXR α and LXR β . Compounds that modulate LXR-mediated transcriptional activation are likely to modulate ABC1 gene expression and thus are useful for modulating HDL cholesterol levels and triglyceride levels. Janowski *et al.* (Proc. Natl. Acad. Sci. USA 96:266-271, 1999) described the role of naturally occurring oxysterols in LXR-dependent transactivation through the promoter for cholesterol 7 α -hydroxylase (Cyp7a), which is the rate limiting enzyme in bile acid synthesis.

Janowski further demonstrated that oxysterols bind directly to LXRs. The position specific mono-oxidation of the sterol side chain is required for LXR high affinity binding and activation. Enhanced binding could be achieved by use of 24-oxo ligands. Oxygens at more than one carbon on the side chain of cholesterol diminished LXR binding and activation as compared to monooxygenated analogs. LXR ligands were found to require a single stereoselective oxygen on the sterol side chain that functioned as a hydrogen acceptor. Introduction of dimethylamide exhibited the greatest binding and activation compared to an ester or carbonyl group.

Compounds known to modulate LXR activity include, without limitation, 24-(S),25-epoxycholesterol; 24(S)-hydroxycholesterol; 22-(R)-hydroxycholesterol; 24(R),25-epoxycholesterol; 22(R)-hydroxy-24(S),25-epoxycholesterol; 22(S)-hydroxy-24(R),25-epoxycholesterol; 24-(S),25-iminocholesterol; methyl-38-hydroxychololate; N,N-dimethyl-3 β -hydroxycholamide; 24(R)-hydroxycholesterol; 22(S)-hydroxycholesterol; 22(R),24(S)-dihydroxycholesterol; 25-hydroxycholesterol; 22(R)-hydroxycholesterol; 22(S)-hydroxycholesterol; 24(S),25-dihydroxycholesterol; 24(R),25-dihydroxycholesterol; 24,25-dehydrocholesterol; 25-epoxy-22(R)-hydroxycholesterol; 20(S)-hydroxycholesterol; (20R,22R)-cholest-5-ene-3 β ,20,22-triol; 4,4-dimethyl-5- α -cholesta-8,14,24-trien-3- β -ol; 7 α -hydroxy-24(S),25-epoxycholesterol; 7 β -hydroxy-24(S),25-epoxycholesterol; 7-oxo-24(S),25-epoxycholesterol; 7 α -hydroxycholesterol; 7-oxocholesterol; and desmosterol. Additional LXR-modulating compounds are described, for example, in Janowski *et al.*, Nature 383:728-731, 1996; Lehman *et al.*, J. Biol. Chem. 272:3137-3140, 1997; and Janowski *et al.*, Proc. Natl. Acad. Sci. 96:266-

271, 1998, each of which is hereby incorporated by reference). In addition one in the art will recognize that synthetic sterols having LXR-modulating activity can be readily identified using screening methods known in the art (see, for example, Janowski *et al.*, Proc. Natl. Acad. Sci. 96:266-271, 1998). Non-steroidal agonists
5 such as RIP140 protein, antibodies (monoclonal or polyclonal) specific for LXR α or LXR β ; tetradecycloxy-furnacarboxylic acid (TOFA); tetradecylthioacetic acid; as well as other fatty acids (see, for example, Tobin *et al.* Molec. Endocrin. 14: 741-752, 2000) are also useful LXR-modulating agents.

Additional transcription factors which may also have an effect in
10 modulating ABC1 gene expression and thereby HDL levels, triglyceride levels, atherosclerosis, and CAD risk include REV-ERB α , SREBP-1 & 2, ADD-1, EBP α , CREB binding protein, P300, HNF 4, RAR, and ROR α .. Exemplary binding sites are depicted in Fig. 3. Additional binding sites for these factors can be found, for example, through examination of the sequence in SEQ ID NO: 1.

15 RXR heterodimerizes with many nuclear receptors, including LXR, and aids in transactivating the target gene. Thus, compounds that modulate RXR-mediated transcriptional activity will also modulate ABC1 expression. Numerous RXR-modulating compounds (rexinoid compounds) are known in the art, including, for example, hetero ethylene derivatives; tricyclic retinoids; trienoic
20 retinoids; benzocycloalkenyl-alka:di- or trienoic acid derivatives; bicyclic-aromatic compounds and their derivatives; bicycylmethyl-aryl acid derivatives; phenyl-methyl heterocyclic compounds; tetrahydro-naphthyl compounds; arylthio-tetrahydro-naphthalene derivatives and heterocyclic analogues; 2,4-pentadienoic acid derivatives; tetralin-based compounds;
25 nonatetraenoic acid derivatives; SR11237; dexamethasone; hydroxy, epoxy, and carboxy derivatives of methoprene; bicyclic benzyl, pyridinyl, thiophene, furanyl,

and pyrrole derivatives; benzofuran-acrylic acid derivatives; aryl-substituted and aryl and (3-oxo-1-propenyl)-substituted benzopyran, benzothiopyran, 1,2-dihydroquinoline, and 5,6-dihydronaphthalene derivatives; vitamin D3 (1,25-dihydroxyvitamin D3) and analogs; 24-hydroxylase inhibitor; mono-or
5 polyenic carboxylic acid derivatives; tetrahydroquinolin-2-one-6 or 7-yl and related derivatives; tetrahydronaphthalene; oxyiminoalkanoic acid derivatives; LG 100268; and LGD 1069. Additional compounds include BRL 49653; troglitazone; pioglitazone; ciglitazone; WAY-120; englitazone; AD 5075; and darglitazone.

PPARs may alter transcription of ABC1 by mechanisms including
10 heterodimerization with retinoid X receptors (RXRs) and then binding to specific proliferator response elements (PPREs). Examples of such PPARs include PPAR α , β , γ and δ . These distinct PPARs have been shown to have transcriptional regulatory effects on different genes. PPAR α is expressed mainly in liver, whereas PPAR γ is expressed in predominantly in adipocytes. Both PPAR α and PPAR γ are found in
15 coronary and carotid artery atherosclerotic plaques and in endothelial cells, smooth muscle cells, monocytes and monocyte-derived macrophages. Activation of PPAR α results in altered lipoprotein metabolism through PPAR α 's effect on genes such as lipoprotein lipase (LPL), apolipoprotein CIII (apo CIII) and apolipoprotein AI (apo AI) and AII (apo AII). PPAR α activation results in overexpression of LPL and
20 apoA-I and apoA-II, but inhibits the expression of apo CIII. PPAR α activation also inhibits inflammation, stimulates lipid oxidation and increases the hepatic uptake and esterification of free fatty acids (FFA's). PPAR α and PPAR γ activation may inhibit nitric oxide (NO) synthase in macrophages and prevent interleukin-1 (IL-1) induced expression of IL-6 and cyclo-oxygenase-2 (COX-2) and thrombin induced
25 endothelin-1 expression secondary to negative transcriptional regulation of NF-KB and activation of protein-1 signaling pathway. It has also been shown that PPAR α

induces apoptosis in monocyte-derived macrophages through the inhibition of NF-KB activity.

Activation of PPAR α can be achieved by compounds such as fibrates, β -estradiol, arachidonic acid derivatives, WY-14,643 and LTB₄ or 8(s)HETE.

5 PPAR γ activation can be achieved through compounds such as thiozolidinedione antidiabetic drugs, 9-HODE and 13-HODE. Additional compounds such as nicotinic acid or HMG CoA reductase inhibitors may also alter the activity of PPARs.

Compounds which alter activity of any of the PPARs (e.g., PPAR α or PPAR γ)

10 may have an effect on ABC1 expression and thereby could affect HDL levels, triglyceride levels, atherosclerosis, and risk of CAD. PPARs are also regulated by fatty acids (including modified fatty acids such as 3 thia fatty acids), leukotrienes such as leukotriene B₄ and prostaglandin J₂, which is a natural activator/ligand for PPAR γ . Drugs that modulate PPARs may therefore have an important effect on

15 modulating lipid levels (including HDL and triglyceride levels) and altering CAD risk. This effect could be achieved through the modulation of ABC1 gene expression. Drugs may also effect ABC1 gene expression and thereby HDL and triglyceride levels, by an indirect effect on PPARs via other transcriptional factors such as adipocyte differentiation and determination factor-1 (ADD-1) and sterol

20 regulatory element binding protein-1 and 2 (SREBP-1 and 2). Drugs with combined PPAR α and PPAR γ agonist activity or PPAR α and PPAR γ agonists given in combination for example, may increase HDL levels or decrease triglyceride levels even more.

A PPAR binding site (PPRE element) is found 5' to the ABC1 gene (Fig. 25 3). Like the PPRE elements found in the C-ACS, HD, CYP4A6 and ApoA-I genes, this PPRE site is a trimer related to the PPRE consensus sequence. Partly

because of its similarity in the number and arrangement of repeats in this PPAR binding site, this element in particular is very likely to be of physiological relevance to the regulation of the ABC1 gene.

5 Additional utility of ABC1 polypeptides, nucleic acids, and modulators

ABC1 may act as a transporter of toxic proteins or protein fragments (*e.g.*, APP) out of cells. Thus, ABC1 agonists/upregulators may be useful in the treatment of other disease areas, including Alzheimer's disease, Niemann-Pick disease, and Huntington's disease.

10 ABC transporters have been shown to increase the uptake of long chain fatty acids from the cytosol to peroxisomes and, moreover, to play a role in β -oxidation of very long chain fatty acids. Importantly, in x-linked adrenoleukodystrophy (ALD), fatty acid metabolism is abnormal, due to defects in the peroxisomal ABC transporter. Any agent that upregulates ABC transporter
15 expression or biological activity may therefore be useful for the treatment of ALD or any other lipid disorder.

ABC1 is expressed in macrophages and is required for engulfment of cells undergoing programmed cell death. The apoptotic process itself, and its regulation, have important implications for disorders such as cancer, one
20 mechanism of which is failure of cells to undergo cell death appropriately. ABC1 may facilitate apoptosis, and as such may represent an intervention point for cancer treatment. Increasing ABC1 expression or activity or otherwise up-regulating ABC1 by any method may constitute a treatment for cancer by increasing apoptosis and thus potentially decreasing the aberrant cellular
25 proliferation characterized by this disease. Conversely, down-regulation of ABC1 by any method may provide opportunity for decreasing apoptosis and allowing

increased proliferation of cells in conditions where cell growth is limited. Such disorders include but are not limited to neurodeficiencies and neurodegeneration, and growth disorders. ABC1 could, therefore, be used as a method for identification of compounds for use in the treatment of cancer, or in the treatment of degenerative disorders.

Agents that have been shown to inhibit ABC1 include, for example, the anti-diabetic agents glibenclamide and glyburide, flufenamic acid, diphenylamine-2-carbonic acid, sulfobromophthalein, and DIDS.

Agents that upregulate ABC1 expression or biological activity include but are not limited to protein kinase A, protein kinase C, vanadate, okadaic acid, and IBMX1.

Those in the art will recognize that other compounds can also modulate ABC1 biological activity, and these compounds are also in the spirit of the invention.

Drug screens based on the ABC1 regulatory regions, gene, or protein

The ABC1 protein and gene can be used in screening assays for identification of compounds which modulate its activity and may be potential drugs to regulate cholesterol or triglyceride levels. The ABC1 5' regulatory sequence and other regulatory regions (e.g., exon 1 and exon 2) can be used in screening assays for identification of compounds which modulate ABC1 expression and may be potential drugs to regulate lipid levels, including, for example, HDL-C, LDL-C, and triglycerides. Drug screens to identify compounds that modulate ABC1 expression may employ an ABC1 regulatory region operably linked to ABC1. Preferably, however, the regulatory region is operably linked to a reporter gene (e.g., a gene encoding GFP, chloramphenicol acetyltransferase, or

beta-galactosidase).

Useful ABC1 proteins include wild-type and mutant ABC1 proteins or protein fragments, in a recombinant form or endogenously expressed. Drug screens to identify compounds acting on the ABC1 expression product may
5 employ any functional feature of the protein. In one example, the phosphorylation state or other post-translational modification is monitored as a measure of ABC1 biological activity. ABC1 has ATP binding sites, and thus assays may wholly or in part test the ability of ABC1 to bind ATP or to exhibit ATPase activity. ABC1, by analogy to similar proteins, is thought to be able to form a channel-like
10 structure; drug screening assays could be based upon assaying for the ability of the protein to form a channel, or upon the ability to transport cholesterol or another molecule, or based upon the ability of other proteins bound by or regulated by ABC1 to form a channel. Alternatively, phospholipid or lipid transport can also be used as measures of ABC1 biological activity.

15 There is evidence that, in addition to its role as a regulator of cholesterol levels, ABC1 also transports anions. Functional assays could be based upon this property, and could employ drug screening technology such as (but not limited to) the ability of various dyes to change color in response to changes in specific ion concentrations in such assays can be performed in vesicles such as liposomes, or
20 adapted to use whole cells.

Drug screening assays can also be based upon the ability of ABC1 or other ABC transporters to interact with other proteins. Such interacting proteins can be identified by a variety of methods known in the art, including, for example, radioimmunoprecipitation, co-immunoprecipitation, co-purification, and yeast
25 two-hybrid screening. Such interactions can be further assayed by means including but not limited to fluorescence polarization or scintillation proximity

methods. Drug screens can also be based upon functions of the ABC1 protein deduced upon X-ray crystallography of the protein and comparison of its 3-D structure to that of proteins with known functions. Such a crystal structure has been determined for the prokaryotic ABC family member HisP, histidine permease. Drug screens can be based upon a function or feature apparent upon creation of a transgenic or knockout mouse, or upon overexpression of the protein or protein fragment in mammalian cells *in vitro*. Moreover, expression of mammalian (e.g., human) ABC1 in yeast or *C. elegans* allows for screening of candidate compounds in wild-type and mutant backgrounds, as well as screens for mutations that enhance or suppress an ABC1-dependent phenotype. Modifier screens can also be performed in ABC1 transgenic or knock-out mice.

Additionally, drug screening assays can also be based upon ABC1 functions deduced upon antisense interference with the gene function. Intracellular localization of ABC1, or effects which occur upon a change in intracellular localization of the protein, can also be used as an assay for drug screening. Immunocytochemical methods will be used to determine the exact location of the ABC1 protein.

Human and rodent ABC1 protein can be used as an antigen to raise antibodies, including monoclonal antibodies. Such antibodies will be useful for a wide variety of purposes, including but not limited to functional studies and the development of drug screening assays and diagnostics. Monitoring the influence of agents (e.g., drugs, compounds) on the expression or biological activity of ABC1 can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase ABC1 gene expression, protein levels, or biological activity can be monitored in clinical trials of subjects exhibiting altered ABC1

gene expression, protein levels, or biological activity. Alternatively, the effectiveness of an agent determined by a screening assay to modulate ABC1 gene expression, protein levels, or biological activity can be monitored in clinical trials of subjects exhibiting decreased altered gene expression, protein levels, or
5 biological activity. In such clinical trials, the expression or activity of ABC1 and, preferably, other genes that have been implicated in, for example, cardiovascular disease can be used to ascertain the effectiveness of a particular drug.

For example, and not by way of limitation, genes, including ABC1, that are modulated in cells by treatment with an agent (e.g., compound, drug or small
10 molecule) that modulates ABC1 biological activity (e.g., identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cholesterol levels, triglyceride levels, or cardiovascular disease, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of ABC1 and other genes implicated in the disorder. The levels of gene
15 expression can be quantified by Northern blot analysis or RT-PCR, or, alternatively, by measuring the amount of protein produced, by one of a number of methods known in the art, or by measuring the levels of biological activity of ABC1 or other genes. In this way, the gene expression can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this
20 response state may be determined before, and at various points during, treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an
agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule,
25 or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject

prior to administration of the agent; (ii) detecting the level of expression of an ABC1 protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the ABC1 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the ABC1 protein, mRNA, or genomic DNA in the pre-administration sample with the ABC1 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of ABC1 to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of ABC1 to lower levels than detected.

The *ABC1* gene or a fragment thereof can be used as a tool to express the protein in an appropriate cell *in vitro* or *in vivo* (gene therapy), or can be cloned into expression vectors which can be used to produce large enough amounts of ABC1 protein to use in *in vitro* assays for drug screening. Expression systems which may be employed include baculovirus, herpes virus, adenovirus, adeno-associated virus, bacterial systems, and eucaryotic systems such as CHO cells. Naked DNA and DNA-liposome complexes can also be used.

Assays of ABC1 activity includes binding to intracellular interacting proteins; interaction with a protein that up-regulates ABC1 activity; interaction with HDL particles or constituents; interaction with other proteins which facilitate interaction with HDL or its constituents; and measurement of cholesterol efflux. Furthermore, assays may be based upon the molecular dynamics of macromolecules, metabolites and ions by means of fluorescent-protein biosensors.

Alternatively, the effect of candidate modulators on expression or activity may be measured at the level of ABC1 protein production using the same general approach in combination with standard immunological detection techniques, such as Western blotting or immunoprecipitation with an ABC1-specific antibody.

5 Again, useful cholesterol- or triglyceride-regulating or anti-CVD therapeutic modulators are identified as those which produce an change in ABC1 polypeptide production. Agonists may also affect ABC1 activity without any effect on expression level.

10 Candidate modulators may be purified (or substantially purified) molecules or may be one component of a mixture of compounds (*e.g.*, an extract or supernatant obtained from cells). In a mixed compound assay, ABC1 expression is tested against progressively smaller subsets of the candidate compound pool (*e.g.*, produced by standard purification techniques, *e.g.*, HPLC or FPLC; Ausubel *et al.*) until a single compound or minimal compound mixture is demonstrated to
15 modulate ABC1 expression.

Agonists, antagonists, or mimetics found to be effective at modulating the level of cellular ABC1 expression or activity may be confirmed as useful in animal models (for example, mice, pigs, rabbits, or chickens). For example, the compound may ameliorate the low HDL levels of mouse or chicken
20 hypoalphalipoproteinemias or may lower the triglyceride levels in animal models.

A compound that promotes an increase in ABC1 expression or activity is considered particularly useful in the invention; such a molecule may be used, for example, as a therapeutic to increase the level or activity of native, cellular ABC1 and thereby treat a low HDL or high triglyceride condition in an animal (for
25 example, a human). If desired, treatment with an agonist of the invention may be combined with any other HDL-raising, triglyceride-lowering, or anti-CVD

therapies.

One method for increasing ABC biological activity is to increase the stabilization of the ABC protein or to prevent its degradation. Thus, it would be useful to identify mutations in an ABC polypeptide (e.g., ABC1) that lead to increased protein stability. These mutations can be incorporated into any protein therapy or gene therapy undertaken for the treatment of low HDL-C or any other condition resulting from loss of ABC1 biological activity. Similarly, compounds that increase the stability of a wild-type ABC polypeptide or decrease its catabolism may also be useful for the treatment of low HDL-C or any other condition resulting from loss of ABC1 biological activity. Such mutations and compounds can be identified using the methods described herein.

In one example, cells expressing an ABC polypeptide having a mutation are transiently metabolically labeled during translation and the half-life of the ABC polypeptide is determined using standard techniques. Mutations that increase the half-life of an ABC polypeptide are ones that increase ABC protein stability. These mutations can then be assessed for ABC biological activity. They can also be used to identify proteins that affect the stability of ABC1 mRNA or protein. One can then assay for compounds that act on these factors or on the ability of these factors to bind ABC1.

In another example, cells expressing wild-type ABC polypeptide are transiently metabolically labeled during translation, contacted with a candidate compounds, and the half-life of the ABC polypeptide is determined using standard techniques. Compounds that increase the half-life of an ABC polypeptide are useful compounds in the present invention.

It is understood that, while ABC1 is the preferred ABC transporter for the drug screens described herein, other ABC transporters can also be used. The

replacement of ABC1 with another ABC transporter is possible because it is likely that ABC transporter family members, such as ABC2, ABCR, or ABC8 will have a similar mechanism of regulation.

Exemplary assays are described in greater detail below.

5

Protein-based assays

ABC1 polypeptide (purified or unpurified) can be used in an assay to determine its ability to bind another protein (including, but not limited to, proteins found to specifically interact with ABC1). The effect of a compound on that binding is then determined.

10

Protein Interaction Assays

ABC1 protein (or a polypeptide fragment thereof or an epitope-tagged form or fragment thereof) is harvested from a suitable source (e.g., from a prokaryotic expression system, eukaryotic cells, a cell-free system, or by immunoprecipitation from ABC1-expressing cells). The ABC1 polypeptide is then bound to a suitable support (e.g., nitrocellulose or an antibody or a metal agarose column in the case of, for example, a his-tagged form of ABC1). Binding to the support is preferably done under conditions that allow proteins associated with ABC1 polypeptide to remain associated with it. Such conditions may include use of buffers that minimize interference with protein-protein interactions. The binding step can be done in the presence and absence of compounds being tested for their ability to interfere with interactions between ABC1 and other molecules. If desired, other proteins (e.g., a cell lysate) are added, and allowed time to associate with the ABC polypeptide. The immobilized ABC1 polypeptide is then washed to remove proteins or other cell constituents that may be non-specifically associated with it

15

20

25

the polypeptide or the support. The immobilized ABC1 polypeptide is then dissociated from its support, and so that proteins bound to it are released (for example, by heating), or, alternatively, associated proteins are released from ABC1 without releasing the ABC1 polypeptide from the support. The released proteins and other cell constituents can be analyzed, for example, by SDS-PAGE gel electrophoresis, Western blotting and detection with specific antibodies, phosphoamino acid analysis, protease digestion, protein sequencing, or isoelectric focusing. Normal and mutant forms of ABC1 can be employed in these assays to gain additional information about which part of ABC1 a given factor is binding to. In addition, when incompletely purified polypeptide is employed, comparison of the normal and mutant forms of the protein can be used to help distinguish true binding proteins.

The foregoing assay can be performed using a purified or semipurified protein or other molecule that is known to interact with ABC1. This assay may include the following steps.

1. Harvest ABC1 protein and couple a suitable fluorescent label to it;
2. Label an interacting protein (or other molecule) with a second, different fluorescent label. Use dyes that will produce different quenching patterns when they are in close proximity to each other versus when they are physically separate (i.e., dyes that quench each other when they are close together but fluoresce when they are not in close proximity);
3. Expose the interacting molecule to the immobilized ABC1 in the presence or absence of a compound being tested for its ability to interfere with an interaction between the two; and
4. Collect fluorescent readout data.

Another assay is includes Fluorescent Resonance Energy Transfer (FRET) assay. This assay can be performed as follows.

1. Provide ABC1 protein or a suitable polypeptide fragment thereof and couple a suitable FRET donor (e.g., nitro-benzoxadiazole (NBD)) to it;
2. Label an interacting protein (or other molecule) with a FRET acceptor (e.g., rhodamine);
3. Expose the acceptor-labeled interacting molecule to the donor-labeled ABC1 in the presence or absence of a compound being tested for its ability to interfere with an interaction between the two; and
4. Measure fluorescence resonance energy transfer.

Quenching and FRET assays are related. Either one can be applied in a given case, depending on which pair of fluorophores is used in the assay.

Membrane permeability assay

The ABC1 protein can also be tested for its effects on membrane permeability. For example, beyond its putative ability to translocate lipids, ABC1 might affect the permeability of membranes to ions. Other related membrane proteins, most notably the cystic fibrosis transmembrane conductance regulator and the sulfonylurea receptor, are associated with and regulate ion channels.

ABC1 or a fragment of ABC1 is incorporated into a synthetic vesicle, or, alternatively, is expressed in a cell and vesicles or other cell sub-structures containing ABC1 are isolated. The ABC1-containing vesicles or cells are loaded with a reporter molecule (such as a fluorescent ion indicator whose fluorescent properties change when it binds a particular ion) that can detect ions (to observe outward movement), or alternatively, the external medium is loaded with such a

molecule (to observe inward movement). A molecule which exhibits differential properties when it is inside the vesicle compared to when it is outside the vesicle is preferred. For example, a molecule that has quenching properties when it is at high concentration but not when it is at another low concentration would be suitable. The movement of the charged molecule (either its ability to move or the kinetics of its movement) in the presence or absence of a compound being tested for its ability to affect this process can be determined.

In another assay, membrane permeability is determined electrophysiologically by measuring ionic influx or efflux mediated by or modulated by ABC1 by standard electrophysiological techniques. A suitable control (e.g., TD cells or a cell line with very low endogenous ABC1 expression) can be used as a control in the assay to determine if the effect observed is specific to cells expressing ABC1.

In still another assay, uptake of radioactive isotopes into or out of a vesicle can be measured. The vesicles are separated from the extravesicular medium and the radioactivity in the vesicles and in the medium is quantitated and compared.

Nucleic acid-based assays

ABC1 nucleic acid may be used in an assay based on the binding of factors necessary for ABC1 gene transcription. The association between the *ABC1* DNA and the binding factor may be assessed by means of any system that discriminates between protein-bound and non-protein-bound DNA (e.g., a gel retardation assay). The effect of a compound on the binding of a factor to *ABC1* DNA is assessed by means of such an assay. In addition to *in vitro* binding assays, *in vivo* assays in which the regulatory regions of the *ABC1* gene are linked to reporter genes can also be performed.

Assays measuring ABC1 stability

A cell-based or cell-free system can be used to screen for compounds based on their effect on the half-life of *ABC1* mRNA or ABC1 protein. The assay may employ labeled mRNA or protein. Alternatively, *ABC1* mRNA may be detected by means of specifically hybridizing probes or a quantitative PCR assay. Protein can be quantitated, for example, by fluorescent antibody-based methods.

In vitro mRNA stability assay

1. Isolate or produce, by *in vitro* transcription, a suitable quantity of ABC1 mRNA;
2. Label the ABC1 mRNA;
3. Expose aliquots of the mRNA to a cell lysate in the presence or absence of a compound being tested for its ability to modulate *ABC1* mRNA stability; and
4. Assess intactness of the remaining mRNA at suitable time points.

In vitro protein stability assay

1. Express a suitable amount of ABC1 protein;
2. Label the protein;
3. Expose aliquots of the labeled protein to a cell lysate in the presence or absence of a compound being tested for its ability to modulate ABC1 protein stability; and
4. Assess intactness of the remaining protein at suitable time points

In vivo mRNA or protein stability assay

1. Incubate cells expressing ABC1 mRNA or protein with a tracer (radiolabeled ribonucleotide or radiolabeled amino acid, respectively) for a very

brief time period (e.g., five minutes) in the presence or absence of a compound being tested for its effect on mRNA or protein stability;

2. Incubate with unlabeled ribonucleotide or amino acid; and

3. Quantitate the ABC1 mRNA or protein radioactivity at time intervals beginning with the start of step 2 and extending to the time when the radioactivity in ABC1 mRNA or protein has declined by approximately 80%. It is preferable to separate the intact or mostly intact mRNA or protein from its radioactive breakdown products by a means such as gel electrophoresis in order to quantitate the mRNA or protein.

Assays measuring inhibition of dominant negative activity

Mutant ABC1 polypeptides are likely to have dominant negative activity (i.e., activity that interferes with wild-type ABC1 function). An assay for a compound that can interfere with such a mutant may be based on any method of quantitating normal ABC1 activity in the presence of the mutant. For example, normal ABC1 facilitates cholesterol efflux, and a dominant negative mutant would interfere with this effect. The ability of a compound to counteract the effect of a dominant negative mutant may be based on cellular cholesterol efflux, or on any other normal activity of the wild-type ABC1 that was inhibitable by the mutant.

Assays measuring phosphorylation

Glu89 in the wild-type chicken ABC1 polypeptide is likely to be part of a phosphorylation motif, and thus elimination of this phosphorylation motif by the E⇒K ABC1 mutation in the WHAM chicken (discussed further below) may be responsible for reduced biological activity of WHAM chicken ABC1. Thus, compounds that modulate the phosphorylation state of ABC1 are likely to be

clinically relevant modulators of human ABC1 activity.

The effect of a compound on ABC1 phosphorylation can be assayed by methods that quantitate phosphates on proteins or that assess the phosphorylation state of a specific residue of a ABC1. Such methods include but are not limited to ³²P labeling and immunoprecipitation, detection with antiphosphoamino acid antibodies (e.g., antiphosphoserine antibodies), phosphoamino acid analysis on 2-dimensional TLC plates, and protease digestion fingerprinting of proteins followed by detection of ³²P-labeled fragments.

Assays measuring other post-translational modifications

The effect of a compound on the post-translational modification of ABC1 is based on any method capable of quantitating that particular modification. For example, effects of compounds on glycosylation may be assayed by treating ABC1 with glycosylase and quantitating the amount and nature of carbohydrate released.

Assays measuring ATP binding

The ability of ABC1 to bind ATP provides another assay to screen for compounds that affect ABC1. ATP binding can be quantitated as follows.

1. Provide ABC1 protein at an appropriate level of purity and reconstitute it in a lipid vesicle;
2. Expose the vesicle to a labeled but non-hydrolyzable ATP analog (such as gamma ³⁵S-ATP) in the presence or absence of compounds being tested for their effect on ATP binding. Note that azido-ATP analogs can be used to allow covalent attachment of the azido-ATP to protein (by means of U.V. light), and permit easier quantitation of the amount of ATP bound to the protein; and

3. Quantitate the amount of ATP analog associated with ABC1

Assays measuring ATPase activity

Quantitation of the ATPase activity of ABC1 can also be assayed for the effect of compounds on ABC1. This is preferably performed in a cell-free assay so as to separate ABC1 from the many other ATPases in the cell. An ATPase assay may be performed in the presence or absence of membranes, and with or without integration of ABC1 protein into a membrane. If performed in a vesicle-based assay, the ATP hydrolysis products produced or the ATP hydrolyzed may be measured within or outside of the vesicles, or both. Such an assay may be based on disappearance of ATP or appearance of ATP hydrolysis products.

For high-throughput screening, a coupled ATPase assay is preferable. For example, a reaction mixture containing pyruvate kinase and lactate dehydrogenase can be used. The mixture includes phosphoenolpyruvate (PEP), nicotinamide adenine dinucleotide (NAD⁺), and ATP. The ATPase activity of ABC1 generates ADP from ATP. The ADP is then converted back to ATP as part of the pyruvate kinase reaction. The product, pyruvate, is then converted to lactate. The latter reaction generates a colored quinone (NADH) from a colorless substrate (NAD⁺), and the entire reaction can be monitored by detection of the color change upon formation of NADH. Since ADP is limiting for the pyruvate kinase reaction, this coupled system precisely monitors the ATPase activity of ABC1.

Assays measuring cholesterol efflux

A transport-based assay can be performed *in vivo* or *in vitro*. For example, the assay may be based on any part of the reverse cholesterol transport process that is readily re-created in culture, such as cholesterol or phospholipid efflux.

Alternatively, the assay may be based on net cholesterol transport in a whole organism, as assessed by means of a labeled substance (such as cholesterol).

For high throughput, fluorescent lipids can be used to measure ABC1-catalyzed lipid efflux. For phospholipids, a fluorescent precursor, C6-NBD-phosphatidic acid, can be used. This lipid is taken up by cells and dephosphorylated by phosphatidic acid phosphohydrolase. The product, NBD-diglyceride, is then a precursor for synthesis of glycerophospholipids like phosphatidylcholine. The efflux of NBD-phosphatidylcholine can be monitored by detecting fluorescence resonance energy transfer (FRET) of the NBD to a suitable acceptor in the cell culture medium. This acceptor can be rhodamine-labeled phosphatidylethanolamine, a phospholipid that is not readily taken up by cells. The use of short-chain precursors obviates the requirement for the phospholipid transfer protein in the media. For cholesterol, NBD-cholesterol ester can be reconstituted into LDL. The LDL can efficiently deliver this lipid to cells via the LDL receptor pathway. The NBD-cholesterol esters are hydrolyzed in the lysosomes, resulting in NBD-cholesterol that can now be transported back to the plasma membrane and efflux from the cell. The efflux can be monitored by the aforementioned FRET assay in which NBD transfers its fluorescence resonance energy to the rhodamine-phosphatidylethanolamine acceptor.

Animal Model Systems

Compounds identified as having activity in any of the above-described assays are subsequently screened in any available animal model system, including, but not limited to, pigs, rabbits, and WHAM chickens. Test compounds are administered to these animals according to standard methods. Test compounds may also be tested in mice bearing mutations in the *ABC1* gene. Additionally,

compounds may be screened for their ability to enhance an interaction between ABC1 and any HDL particle constituent such as ApoAI, ApoAII, or ApoE.

The cholesterol efflux assay as a drug screen

5 The cholesterol efflux assay measures the ability of cells to transfer cholesterol to an extracellular acceptor molecule and is dependent on ABC1 function. In this procedure, cells are loaded with radiolabeled cholesterol by any of several biochemical pathways (Marcil *et al.*, Arterioscler. Thromb. Vasc. Biol. 19:159-169, 1999). Cholesterol efflux is then measured after incubation for
10 various times (typically 0 to 24 hours) in the presence of HDL3 or purified ApoAI. Cholesterol efflux is determined as the percentage of total cholesterol in the culture medium after various times of incubation. ABC1 expression levels and/or biological activity are associated with increased efflux while decreased levels of ABC1 are associated with decreased cholesterol efflux.

15 This assay can be readily adapted to the format used for drug screening, which may consist of a multi-well (*e.g.*, 96-well) format. Modification of the assay to optimize it for drug screening would include scaling down and streamlining the procedure, modifying the labeling method, using a different cholesterol acceptor, altering the incubation time, and changing the method of
20 calculating cholesterol efflux. In all these cases, the cholesterol efflux assay remains conceptually the same, though experimental modifications may be made. A transgenic mouse overexpressing ABC1 would be expected to have higher than normal HDL levels.

Knock-out mouse model

25 An animal, such as a mouse, that has had one or both ABC1 alleles

inactivated (*e.g.*, by homologous recombination) is likely to have low HDL-C levels and higher than normal triglyceride levels, and thus is a preferred animal model for screening for compounds that raise HDL-C levels or lower triglyceride levels. Such an animal can be produced using standard techniques. In addition to the initial screening of test compounds, the animals having mutant ABC1 genes are useful for further testing of efficacy and safety of drugs or agents first identified using one of the other screening methods described herein. Cells taken from the animal and placed in culture can also be exposed to test compounds. HDL-C and triglyceride levels can be measured using standard techniques, such as those described herein.

WHAM chickens: an animal model for low HDL cholesterol

Wisconsin Hypo-Alpha Mutant (WHAM) chickens arose by spontaneous mutation in a closed flock. Mutant chickens came to attention through their a Z-linked white shank and white beak phenotype referred to as ‘recessive white skin’ (McGibbon, 1981) and were subsequently found to have a profound deficiency of HDL (Poernama *et al.*, 1990).

This chicken low HDL locus (Y) is Z-linked, or sex-linked. (In birds, females are ZW and males are ZZ). Genetic mapping placed the Y locus on the long arm of the Z chromosome (Bitgood, 1985), proximal to the ID locus (Bitgood, 1988). Examination of current public mapping data for the chicken genome mapping project, ChickMap (maintained by the Roslin Institute; <http://www.ri.bbsrc.ac.uk/chickmap/ChickMapHomePage.html>) showed that a region of synteny with human chromosome 9 lies on the long arm of the chicken Z chromosome (Zq) proximal to the ID locus. Evidence for this region of synteny is the location of the chicken aldolase B locus (ALDOB) within this region. The

human ALDOB locus maps to chromosome 9q22.3 (The Genome Database, <http://gdbwww.gdb.org/>), not far from the location of human ABC1. This comparison of maps showed that the chicken Zq region near chicken ALDOB and the human 9q region near human ALDOB represent a region of synteny between human and chicken.

Since a low HDL locus maps to the 9q location in humans and to the Zq region in chickens, these low HDL loci are most probably located within the syntenic region. Thus we predicted that ABC1 is mutated in WHAM chickens. In support of this, we have previously identified an E⇒K mutation at a position that corresponds to amino acid 89 of human ABC1. This non-conservative substitution is at a position that is conserved among human, mouse, and chicken, indicating that it is in a region of the protein likely to be of functional importance.

Discovery of the WHAM mutation in the amino-terminal portion of the ABC1 protein also establishes the importance of the amino-terminal region. This region may be critical because of association with other proteins required to carry out cholesterol efflux or related tasks. It may be an important regulatory region (there is a phosphorylation site for casein kinase near the mutated residue), or it may help to dictate a precise topological relationship with cellular membranes (the N-terminal 60 amino acid region contains a putative membrane-spanning or membrane-associated segment).

The amino-terminal region of the protein (up to the first 6-TM region at approximately amino acid 639) is an ideal tool for screening factors that affect ABC1 activity. It can be expressed as a truncated protein in ABC1 wild-type cells in order to test for interference of the normal ABC1 function by the truncated protein. If the fragment acts in a dominant negative way, it could be used in immunoprecipitations to identify proteins that it may be competing away from the

normal endogenous protein.

The C-terminus also lends itself to such experiments, as do the intracellular portions of the molecule, expressed as fragments or tagged or fusion proteins, in the absence of transmembrane regions.

5 Since it is possible that there are several genes in the human genome which affect cholesterol efflux, it is important to establish that any animal model to be used for a human genetic disease represents the homologous locus in that animal, and not a different locus with a similar function. The evidence above establishes that the chicken Y locus and the human chromosome 9 low HDL locus are
10 homologous. WHAM chickens are therefore an important animal model for the identification and testing of drugs that modulate cholesterol efflux.

The WHAM chickens' HDL deficiency syndrome is not, however, known to be associated with an increased susceptibility to atherosclerosis in chickens. This may reflect the shorter lifespan or, more likely, the impaired absorption of
15 dietary cholesterol in these chickens. We propose the WHAM chicken as a model for human low HDL for the development and testing of drugs to raise HDL in humans. Such a model could be employed in several forms, through the use of cells or other derivatives of these chickens, or by the use of the chickens themselves in tests of drug effectiveness, toxicity, and other drug development
20 purposes.

Therapy

Compounds of the invention, including but not limited to, ABC1 polypeptides, *ABC1* nucleic acids, other ABC transporters, LXR-modulating
25 compounds, RXR-modulating compounds, and any therapeutic agent that modulates biological activity or expression of ABC1 identified using any of the

methods disclosed herein, may be administered with a pharmaceutically-
acceptable diluent, carrier, or excipient, in unit dosage form. Conventional
pharmaceutical practice may be employed to provide suitable formulations or
compositions to administer such compositions to patients. Any appropriate route
5 of administration may be employed, for example, intravenous, perenteral,
subcutaneous, intramuscular, intracranial, intraorbital, ophthalmic, intraventricular,
intracapsular, intraspinal, intracisternal, intraperitoneal, intranasal, aerosol, or oral
administration. Therapeutic formulations may be in the form of liquid solutions or
suspension; for oral administration, formulations may be in the form of tablets or
10 capsules; and for intranasal formulations, in the form of powders, nasal drops, or
aerosols.

Methods well known in the art for making formulations are found in, for
example, Remington: The Science and Practice of Pharmacy, (19th ed.) ed. A.R.
Gennaro AR., 1995, Mack Publishing Company, Easton, PA. Formulations for
15 parenteral administration may, for example, contain excipients, sterile water, or
saline, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin,
or hydrogenated naphthalenes. Biocompatible, biodegradable lactide polymer,
lactide/glycolide copolymer, or polyoxyethylene-polyoxypropylene copolymers
may be used to control the release of the compounds. Other potentially useful
20 parenteral delivery systems for agonists of the invention include ethylenevinyl
acetate copolymer particles, osmotic pumps, implantable infusion systems, and
liposomes. Formulations for inhalation may contain excipients, or example,
lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-
lauryl ether, glycocholate and deoxycholate, or may be oily solutions for
25 administration in the form of nasal drops, or as a gel.

Compounds

In general, novel drugs for the treatment of aberrant lipid levels and/or CVD are identified from large libraries of both natural product or synthetic (or semi-synthetic) extracts or chemical libraries according to methods known in the art. Those skilled in the field of drug discovery and development will understand that the precise source of test extracts or compounds is not critical to the screening procedure(s) of the invention. Accordingly, virtually any number of chemical extracts or compounds can be screened using the exemplary methods described herein. Examples of such extracts or compounds include, but are not limited to, plant-, fungal-, prokaryotic- or animal-based extracts, fermentation broths, and synthetic compounds, as well as modification of existing compounds. Numerous methods are also available for generating random or directed synthesis (*e.g.*, semi-synthesis or total synthesis) of any number of chemical compounds, including, but not limited to, saccharide-, lipid-, peptide-, and nucleic acid-based compounds. Synthetic compound libraries are commercially available from Brandon Associates (Merrimack, NH) and Aldrich Chemical (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant, and animal extracts are commercially available from a number of sources, including Biotics (Sussex, UK), Xenova (Slough, UK), Harbor Branch Oceanographics Institute (Ft. Pierce, FL), and PharmaMar, U.S.A. (Cambridge, MA). In addition, natural and synthetically produced libraries are produced, if desired, according to methods known in the art, *e.g.*, by standard extraction and fractionation methods. Furthermore, if desired, any library or compound is readily modified using standard chemical, physical, or biochemical methods.

In addition, those skilled in the art of drug discovery and development readily understand that methods for dereplication (*e.g.*, taxonomic dereplication,

biological dereplication, and chemical dereplication, or any combination thereof) or the elimination of replicates or repeats of materials already known for their HDL-raising, triglyceride-lowing, or anti-CVD activities, or known for their ability to modulate ABC1 gene expression or ABC1 biological activity should be employed whenever possible.

When a crude extract is found to modulate ABC1 gene expression, ABC1 biological activity, or a combination thereof, further fractionation of the positive lead extract is necessary to isolate chemical constituent responsible for the observed effect. Thus, the goal of the extraction, fractionation, and purification process is the careful characterization and identification of a chemical entity within the crude extract having HDL-raising, triglyceride-lowering, or anti-CVD activities, ability to modulate ABC1 gene expression, or a combination thereof. The same *in vivo* and *in vitro* assays described herein for the detection of activities in mixtures of compounds can be used to purify the active component and to test derivatives thereof. Methods of fractionation and purification of such heterogeneous extracts are known in the art. If desired, compounds shown to be useful agents for the treatment of pathogenicity are chemically modified according to methods known in the art. Compounds identified as being of therapeutic value are subsequently analyzed using any standard animal model of diabetes or obesity known in the art.

It is understood that compounds that modulate activity of proteins that modulate ABC1 gene expression or activity are useful compounds for modulating HDL-C levels and triglyceride levels. Exemplary compounds are provided herein; others are known in the art.

Compounds that are structurally related to cholesterol, or that mimic ApoAI or a related apolipoprotein, and increase ABC1 biological activity are particularly

useful compounds in the invention. Other compounds, known to act on the MDR protein, can also be used or derivatized and assayed for their ability to increase ABC1 biological activity. Exemplary MDR modulators are PSC833, bromocriptine, and cyclosporin A. Other examples of compounds that may be assayed for the ability to increase ABC1 biological activity include oxysterols and their derivatives.

Screening patients having low HDL-C or high triglyceride levels

ABC1 expression, biological activity, and mutational analysis can each serve as a diagnostic tool for low HDL or higher than normal triglyceride levels; thus determination of the genetic subtyping of the *ABC1* gene sequence can be used to subtype low HDL or higher than normal triglyceride individuals or families to determine whether the low HDL or higher than normal triglyceride phenotype is related to ABC1 function. This diagnostic process can lead to the tailoring of drug treatments according to patient genotype, including prediction of side-effects upon administration of HDL increasing or triglyceride lowering drugs (referred to herein as pharmacogenomics). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual is examined to determine the ability of the individual to respond to a particular agent).

Agents, or modulators which have a stimulatory or inhibitory effect on ABC1 biological activity or gene expression can be administered to individuals to treat disorders (e.g., cardiovascular disease, low HDL cholesterol, or a higher than normal triglyceride level) associated with aberrant ABC1 activity. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign

compound or drug) of the individual may be considered. Differences in efficacy of
therapeutics can lead to severe toxicity or therapeutic failure by altering the
relation between dose and blood concentration of the pharmacologically active
drug. Thus, the pharmacogenomics of the individual permits the selection of
5 effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a
consideration of the individual's genotype. Such pharmacogenomics can further
be used to determine appropriate dosages and therapeutic regimens. Accordingly,
the activity of ABC1 protein, expression of ABC1 nucleic acid, or mutation
content of ABC1 genes in an individual can be determined to thereby select
10 appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in
the response to drugs due to altered drug disposition and abnormal action in
affected persons (Eichelbaum, M., Clin. Exp. Pharmacol. Physiol., 23:983-985,
1996; Linder, M. W., Clin. Chem., 43:254-266, 1997). In general, two types of
15 pharmacogenetic conditions can be differentiated. Genetic conditions transmitted
as a single factor altering the way drugs act on the body (altered drug action) or
genetic conditions transmitted as single factors altering the way the body acts on
drugs (altered drug metabolism). Altered drug action may occur in a patient
having a polymorphism (e.g., an single nucleotide polymorphism or SNP) in
20 promoter, intronic, or exonic sequences of ABC1. Thus by determining the
presence and prevalence of polymorphisms allow for prediction of a patient's
response to a particular therapeutic agent. In particular, polymorphisms in the
promoter region may be critical in determining the risk of HDL deficiency, higher
than normal triglyceride level, and CVD.

25 In addition to the mutations in the *ABC1* gene described herein, we have
detected polymorphisms in the human *ABC1* gene (Fig. 4). These polymorphisms

are located in promoter, intronic, and exonic sequence of ABC1. Using standard methods, such as direct sequencing, PCR, SSCP, or any other polymorphism-detection system, one could easily ascertain whether these polymorphisms are present in a patient prior to the establishment of a drug treatment regimen for a patient having low HDL, a higher than normal triglyceride level, cardiovascular disease, or any other ABC1-mediated condition. It is possible that some these polymorphisms are, in fact, weak mutations. Individuals harboring such mutations may have an increased risk for cardiovascular disease; thus, these polymorphisms may also be useful in diagnostic assays.

Other Embodiments

All publications mentioned in this specification are herein incorporated by reference to the same extent as if each independent publication was specifically and individually indicated to be incorporated by reference.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications. This application is intended to cover any variations, uses, or adaptations following, in general, the principles of the invention and including such departures from the present disclosure within known or customary practice within the art to which the invention pertains and may be applied to the essential features hereinbefore set forth.

What is claimed is:

1. A method for treating a patient diagnosed as having a lower than normal HDL-cholesterol level, a higher than normal triglyceride level, or a cardiovascular disease, said method comprising administering to said patient a compound that modulates LXR-mediated transcriptional activity.

5

2. The method of claim 1, wherein said compound is an oxysterol.

3. The method of claim 1, wherein said compound is selected from the group consisting of 24-(S),25-epoxycholesterol; 24(S)-hydroxycholesterol;

10

22-(R)-hydroxycholesterol; 24(R),25-epoxycholesterol;

22(R)-hydroxy-24(S),25-epoxycholesterol;

22(S)-hydroxy-24(R),25-epoxycholesterol; 24-(S),25-iminocholesterol;

methyl-38-hydroxycholesterol; N,N-dimethyl-3 β -hydroxycholesterol;

24(R)-hydroxycholesterol; 22(S)-hydroxycholesterol;

15

22(R),24(S)-dihydroxycholesterol; 25-hydroxycholesterol;

22(R)-hydroxycholesterol; 22(S)-hydroxycholesterol;

24(S),25-dihydroxycholesterol; 24(R),25-dihydroxycholesterol;

24,25-dehydrocholesterol; 25-epoxy-22(R)-hydroxycholesterol;

20(S)-hydroxycholesterol; (20R,22R)-cholest-5-ene-3 β ,20,22-triol;

20

4,4-dimethyl-5- α -cholesta-8,14,24-trien-3- β -ol;

7 α -hydroxy-24(S),25-epoxycholesterol; 7 β -hydroxy-24(S),25-epoxycholesterol;

7-oxo-24(S),25-epoxycholesterol; 7 α -hydroxycholesterol; 7-oxocholesterol;

and desmosterol.

25

4. The method of claim 1, wherein said LXR is LXR α .

5. A method for treating a patient diagnosed as having a lower than normal HDL-cholesterol level, a higher than normal triglyceride level, or a cardiovascular disease, said method comprising administering to said patient a compound that modulates RXR-mediated transcriptional activity.

6. The method of claim 5, wherein said RXR is RXR α .

7. A method for determining whether a candidate compound modulates ABC1 expression, said method comprising the steps of:

(a) providing a nucleic acid molecule comprising an ABC1 regulatory region or promoter linked to a reporter gene;

(b) contacting said nucleic acid molecule with said candidate compound; and

(c) measuring expression of said reporter gene, wherein altered reporter gene expression, relative to said reporter gene expression of a corresponding control nucleic acid molecule not contacted with said compound, indicates that said candidate compound modulates ABC1 expression.

8. The method of claim 7, wherein said regulatory region comprises 50 consecutive nucleotides selected from nucleotides 5854 to 6694, 7756 to 8318, 10479 to 10825, 15214 to 16068, 21636 to 22111, 27898 to 28721, 32951 to 33743, 36065 to 36847, 39730 to 40577, 4543 to 5287, and 45081 to 55639 of SEQ ID NO: 1.

9. The method of claim 7, wherein said regulatory region comprises a binding site for a transcription factor selected from a group consisting of LXRs, RXRs, RORs, SREBPs, and PPARs.

5 10. A substantially pure nucleic acid molecule comprising a region that is substantially identical to at least fifty contiguous nucleotides of nucleotides 5854 to 6694, 7756 to 8318, 10479 to 10825, 15214 to 16068, 21636 to 22111, 27898 to 28721, 32951 to 33743, 36065 to 36847, 39730 to 40577, 4543 to 5287, or 45081 to 55639 of SEQ ID NO: 1.

10 11. A substantially pure nucleic acid molecule comprising a region that is substantially identical to nucleotides 1 to 28,707 of SEQ ID NO: 1.

15 12. A substantially pure nucleic acid molecule comprising a region that is substantially identical to nucleotides 29,011 to 53,228 of SEQ ID NO: 1.

13. A cell expressing the nucleic acid molecule of claim 10.

20 14. A non-human mammal expressing the nucleic acid molecule of claim 10.

25 15. A method of treating a human having a higher than normal triglyceride level, said method comprising administering to said human an ABC1 polypeptide, or triglyceride-regulating fragment thereof.

16. The method of claim 15, wherein said ABC1 polypeptide has the

sequence of SEQ ID NO: 5.

17. The method of claim 15, wherein said ABC1 polypeptide comprises a R⇒K mutation at position 219 or a V⇒A mutation at position 399.

5

18. The method of claim 15, wherein said ABC1 polypeptide comprises a mutation that increases its stability.

19. The method of claim 15, wherein said ABC1 polypeptide comprises a mutation that increases its biological activity.

10

20. A method of treating a human having a higher than normal triglyceride level, said method comprising administering to said human a nucleic acid molecule encoding an ABC1 polypeptide or a triglyceride-regulating fragment thereof.

15

21. The method of claim 20, wherein said ABC1 polypeptide has the amino acid sequence of SEQ ID NO: 5.

22. The method of claim 20, wherein said ABC1 polypeptide comprises a R⇒K mutation at position 219 or a V⇒A mutation at position 399.

20

23. The method of claim 20, wherein said ABC1 polypeptide comprises a mutation that increases its stability.

24. The method of claim 20, wherein said ABC1 polypeptide comprises a mutation that increases its biological activity.

25

25. The method of claim 20, wherein said biological activity is regulation of cholesterol.

26. The method of claim 20, wherein said human has a lower than normal HDL-cholesterol level.

27. A method of treating a human having a higher than normal triglyceride level, said method comprising administering to said human a compound that increases ABC1 biological activity or that mimics the activity of wild-type ABC1, R219K ABC1, or V399A ABC1.

28. A non-human mammal comprising a transgene comprising a nucleic acid molecule encoding a dominant-negative ABC1 polypeptide, said dominant-negative polypeptide comprising a M \Rightarrow T mutation at position 1091.

29. A method for determining whether a candidate compound decreases the inhibition of a dominant-negative ABC1 polypeptide, said dominant-negative polypeptide comprising a M \Rightarrow T mutation at position 1091, said method comprising the steps of:

- (a) providing a cell expressing a dominant-negative ABC1 polypeptide;
- (b) contacting said cell with said candidate compound; and
- (c) measuring ABC1 biological activity of said cell,

wherein an increase in said ABC1 biological activity, relative to said ABC1 biological activity in a corresponding control cell not contacted with said compound, indicates that said candidate compound decreases the inhibition of a dominant-negative ABC1 polypeptide.

30. A method for predicting a person's response to a triglyceride-lowering drug, comprising determining whether the person has a polymorphism in an ABC1 gene, promoter, or regulatory sequence that alters the person's response to said drug.

5

31. A method for determining whether a candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

(a) providing a chicken comprising a mutation in an *ABC1* gene;

(b) administering said candidate compound to said chicken; and

10

(c) measuring ABC1 biological activity in said chicken,

wherein altered ABC1 biological activity, relative to said ABC1 biological activity in a corresponding control chicken not contacted with said compound, indicates that said candidate compound is useful for modulating triglyceride levels.

15

32. The method of claim 31, wherein said ABC1 biological activity is transport of cholesterol.

33. A method for determining whether a candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

20

(a) providing a cell expressing an ABC1 polypeptide comprising amino acids 1 to 60 of SEQ ID NO: 5;

(b) contacting said cell with said candidate compound; and

(c) measuring ABC1 biological activity of said cell,

25

wherein altered ABC1 biological activity, relative to said ABC1 biological activity in a corresponding control cell not contacted with said compound, indicates that said candidate compound is useful for modulating triglyceride levels.

34. A method for determining whether a candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

- (a) providing a cell expressing an *ABC1* gene or a fragment thereof;
- (b) contacting said cell with said candidate compound; and
- 5 (c) measuring ABC1 expression of said cell,

wherein altered ABC1 expression, relative to said ABC1 expression in a corresponding control cell not contacted with said candidate compound, indicates that said candidate compound is useful for modulating triglyceride levels.

10 35. A method for determining whether a candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

(a) providing an ABC1 polypeptide comprising amino acids 1 to 60 of SEQ ID NO: 1;

(b) contacting said polypeptide with said candidate compound; and

15 (c) measuring ABC1 biological activity, wherein a change in ABC1 biological activity, relative to said ABC1 biological activity of a corresponding control ABC1 polypeptide not contacted with said compound, indicates that said candidate compound is useful for modulating triglyceride levels.

20 36. A method for determining whether a candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

(a) providing an ABC1 polypeptide comprising amino acids 1 to 60 of SEQ ID NO: 5;

(b) contacting said polypeptide with said candidate compound; and

25 (c) measuring expression of said ABC1 polypeptide,
wherein a change in expression of said ABC1 polypeptide, relative to said

expression of a corresponding control ABC1 polypeptide not contacted with said compound, indicates that said candidate compound is useful for modulating triglyceride levels.

5 37. A method for determining whether candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

 (a) providing an ABC1 polypeptide comprising amino acids 1 to 60 of SEQ ID NO: 5;

 (b) contacting said polypeptide with said candidate compound; and

10 (c) measuring binding of said ABC1 polypeptide to said candidate compound, wherein binding of said ABC1 polypeptide to said compound indicates that said candidate compound is useful for modulating triglyceride levels.

 38. A method for determining whether candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

 (a) providing (i) an ABC1 polypeptide comprising amino acids 1 to 60 of SEQ ID NO: 5, and (ii) a second polypeptide that interacts with said ABC1 polypeptide;

 (b) contacting said polypeptides with said candidate compound; and

20 (c) measuring interaction of said ABC1 polypeptide with said second polypeptide, wherein an alteration in the interaction of said ABC1 polypeptide with said second polypeptide indicates that said candidate compound is useful for modulating triglyceride levels.

25 39. A method for determining whether a candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

(a) providing a cell comprising an ABC1 polypeptide comprising amino acids 1 to 60 of SEQ ID NO: 5;

(b) contacting said cell with said candidate compound; and

(c) measuring the half-life of said ABC1 polypeptide,

5 wherein an increase in said half-life, relative to said half-life in a corresponding control cell not contacted with said compound, indicates that said candidate compound is useful for modulating triglyceride levels.

10 40. A method for determining whether a candidate compound is useful for modulating triglyceride levels, said method comprising the steps of:

(a) providing an ABC1 polypeptide in a lipid membrane;

(b) contacting said polypeptide with said candidate compound; and

(c) measuring ABC1-mediated lipid transport across said lipid membrane,

15 wherein a change in lipid transport, relative to said lipid transport of a corresponding control ABC1 polypeptide not contacted with said compound, indicates that said candidate compound is useful for modulating triglyceride levels.

20 41. The method of claim 35-38, or 40, wherein said ABC1 polypeptide is in a cell-free system.

42. The method of claim 35-38, or 40, wherein said ABC1 polypeptide is in a cell.

25 43. The method of claim 42, wherein said cell is from a WHAM chicken.

44. The method of claim 42, wherein said cell is in a human or in a non-

human mammal.

45. The method of claim 44, wherein said animal is a WHAM chicken.

5 46. The method of claim 31, 33, or 35, wherein said biological activity is transport of lipid or interleukin-1.

47. The method of claim 46, wherein said lipid is cholesterol.

10 48. The method of claim 47, wherein said cholesterol is HDL-cholesterol.

49. The method of claim 31, 33, or 35, wherein said biological activity is binding or hydrolysis of ATP by the ABC1 polypeptide.

15 50. A method of determining a propensity for a disease or condition in a subject, wherein said disease or condition is selected from the group consisting of a lower than normal HDL level, a higher than normal triglyceride level, and a cardiovascular disease, said method comprising determining the presence or absence of at least one ABC1 polymorphism in the polynucleotide sequence of an ABC1 regulatory region, promoter, or coding sequence or in the amino acid
20 sequence of an ABC1 protein in a sample obtained from said subject, wherein the presence or absence of said ABC1 polymorphism is indicative of a risk for said disease or condition.

25 51. The method of claim 50, further comprising analyzing at least five ABC1 polymorphic sites in said polynucleotide sequence or said amino acid

sequence.

52. A method for determining whether an ABC1 polymorphism is indicative of a risk for a disease or condition in a subject, wherein said disease or condition is selected from the group consisting of a lower than normal HDL level, a higher than normal triglyceride level, and a cardiovascular disease, said method comprising the steps of:

(a) determining whether the prevalence of said disease or condition in a first subject or set of subjects differs from said prevalence of said disease or condition in a second subject or set of subjects;

(b) analyzing the polynucleotide sequence of an ABC1 regulatory region, promoter, or coding sequence or the amino acid sequence of an ABC1 protein in a sample obtained from said first subject or set of subjects and said second subject or set of subjects; and

(c) determining whether at least one ABC1 polymorphism differs between said first subject or set of subjects and said second subject or set of subjects, wherein the presence or absence of said ABC1 polymorphism is correlated with said prevalence of said disease or condition, thereby determining whether said ABC1 polymorphism is indicative of said risk.

53. The method of claim 52, further comprising analyzing at least five ABC1 polymorphic sites in said polynucleotide sequence or said amino acid sequence.

54. An electronic database comprising a plurality of sequence records of ABC1 polymorphisms correlated to records of predisposition to or prevalence of a disease or condition selected from the group consisting of a lower than normal

HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

55. A method for selecting a preferred therapy for modulating ABC1 activity or expression in a subject, said method comprising:

(a) determining the presence or absence of at least one ABC1 polymorphism in the polynucleotide sequence of an ABC1 regulatory region, promoter, or coding sequence or in the amino acid sequence of an ABC1 protein in a sample obtained from said subject, wherein the presence or absence of said ABC1 polymorphism is indicative of the safety or efficacy of at least one therapy for modulating ABC1 expression or activity; and

(b) determining a preferred therapy for modulating ABC1 expression or activity in said subject.

56. The method of claim 55, further comprising analyzing at least five ABC1 polymorphic sites in said polynucleotide sequence or said amino acid sequence.

57. A method for determining whether a candidate compound is useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease; said method comprising the steps of:

(a) providing an assay system having a measurable ABC1 biological activity;

(b) contacting said assay system with said candidate compound; and

(c) measuring ABC1 biological activity or ABC1 phosphorylation, wherein modulation of ABC1 biological activity or ABC1 phosphorylation,

relative to said ABC1 biological activity or ABC1 phosphorylation in a corresponding control assay system not contacted with said candidate compound, indicates that said candidate compound is useful for the treatment of said disease or condition.

5

58. The method of claim 57, wherein said assay system is a cell based system

10

59. The method of claim 57, wherein said assay system is a cell free system.

15

60. A method for identifying a compound to be tested for an ability to ameliorate a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, said method comprising the steps of:

20

(a) contacting a subject or cell with a candidate compound;
(b) measuring ABC1 expression, activity, or protein phosphorylation in said subject or cell; wherein altered ABC1 expression, activity, or protein phosphorylation; relative to said ABC1 expression, activity, or protein phosphorylation in a corresponding control subject or cell not contacted with said candidate compound; identifies said candidate compound as a compound to be tested for an ability to ameliorate said disease or condition.

25

61. The method of claim 57 or 60, wherein said candidate compound modulates said ABC1 protein phosphorylation and said ABC1 activity.

62. A method for determining whether a candidate compound is useful for

modulating a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, said method comprising the steps of:

- (a) providing a cell expressing an ABC1 gene or a fragment thereof;
- (b) contacting said cell with said candidate compound; and
- (c) measuring ABC1 activity of said cell, wherein altered ABC1 activity, relative to said ABC1 activity in a corresponding control cell not contacted with said compound, indicates that said candidate compound is useful for modulating said disease or condition.

63. A method for determining whether a candidate compound is useful for modulating a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, said method comprising the steps of:

- (a) contacting a cell expressing an ABC1 protein with said candidate compound;
- (b) measuring the phosphorylation of said ABC1 protein; wherein altered ABC1 protein phosphorylation, relative to said ABC1 protein phosphorylation in a corresponding control cell not contacted with said candidate compound, indicates that said is useful for modulating said disease or condition.

64. A compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, wherein said compound modulates ABC1 biological activity, and wherein said compound is identified by the steps of:

- (a) providing an assay system having a measurable ABC1 biological

activity;

(b) contacting said assay system with said compound; and

(c) measuring ABC1 biological activity, wherein modulation of ABC1 biological activity, relative to said ABC1 biological activity in a corresponding control assay system not contacted with said compound, indicates that said compound is useful for the treatment of said disease or condition.

65. A compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, wherein said compound induces a change in ABC1 biological activity that mimics the change in ABC1 biological activity induced by the R219K ABC1 mutation.

66. A compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, wherein said compound binds or interacts with residue R219 of ABC1, thereby mimicking the change in ABC1 activity induced by the R219K ABC1 mutation.

67. A compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, wherein said compound induces a change in ABC1 biological activity that mimics the change in ABC1 biological activity induced by the V339A ABC1 mutation.

68. A compound useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher

than normal triglyceride level, and a cardiovascular disease, wherein said compound binds or interacts with residue V399 of ABC1, thereby mimicking the change in ABC1 activity induced by the V399A ABC1 mutation.

5 69. A compound that modulates ABC1 activity and binds or interacts with an amino acid of ABC1, wherein said amino acid is a residue selected from amino acids 119 to 319 of ABC1 (SEQ ID NO: 5) or amino acids 299 to 499 of ABC1 (SEQ ID NO: 5).

10 70. A method for determining whether a candidate compound is useful for the treatment a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, said method comprising the steps of:

- 15 (a) providing an assay system having a measurable LXR biological activity;
 (b) contacting the assay system with said candidate compound; and
 (c) measuring LXR biological activity, wherein modulation of LXR biological activity, relative to said LXR biological activity in a corresponding control assay system not contacted with said candidate compound, indicates that said candidate compound is useful for the treatment of said disease or condition.

20 71. A method for determining whether a candidate compound is useful for modulating ABC1 biological activity, said method comprising the steps of:

- 25 (a) providing an assay system having a measurable LXR biological activity;
 (b) contacting said assay system with said candidate compound; and
 (c) measuring LXR biological activity, wherein modulation of LXR biological activity, relative to said LXR biological activity in a corresponding control assay system not contacted with said candidate compound, indicates that

said candidate compound is useful for modulating ABC1 biological activity.

72. The method of claim 71, wherein said LXR biological activity is modulation of ABC1 expression.

5

73. A method for identifying a compound to be tested for an ability to modulate ABC1 biological activity, said method comprising the steps of:

(a) contacting a subject or cell with a candidate compound;

(b) assaying the activity of the LXR gene product in said subject or cell;

10

wherein modulation of said activity, relative to said activity in a corresponding control subject or cell not contacted with said candidate compound, identifies said candidate compound as a compound to be tested for an ability to modulate the biological activity of ABC1.

15

74. Use of an LXR gene product in an assay to identify compounds useful for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

20

75. Use of a compound that modulates the activity or expression of an LXR gene product for the treatment of a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

25

76. A method for identifying a compound to be tested for an ability to treat a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a

cardiovascular disease, said method comprising the steps of:

- (a) providing an assay system having a measurable LXR biological activity;
- (b) contacting said assay system with the candidate compound; and
- (c) measuring LXR biological activity, wherein modulation of said LXR biological activity, relative to said LXR biological activity in a corresponding control assay system not contacted with said candidate compound, identifies said candidate compound as a compound to be tested for an ability to treat said disease or condition.

77. A method for screening an candidate LXR agonist for the ability to treat a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, said method comprising the steps of:

- (a) contacting said a with said candidate LXR agonist; and
- (b) measuring cholesterol efflux activity of said cell, wherein an increase in said cholesterol efflux activity in said cell, relative to said cholesterol efflux in a corresponding control cell not contacted with said candidate LXR agonist, indicates that said candidate LXR agonist is useful for treating said disease or condition.

78. A method for screening a candidate LXR modulating compound for the ability to treat a disease or condition selected from the group consisting of a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease, said method comprising the steps of:

- (a) contacting a cell with said candidate LXR modulating compound; and
- (b) measuring ABC1 biological activity of said cell; wherein an increase in ABC1 biological activity in said cell, relative to said ABC1 biological activity in a

corresponding control cell not contacted with said LXR modulating compound, indicates that said LXR modulating compound is useful for treating said disease or condition.

5 79. The method of any one of claims 71-78, wherein said cell or assay system comprises an exogenously supplied copy of an LXRE selected from the group consisting of SEQ ID NO: 94, SEQ ID NO: 92, and the LXRE consensus motif at nucleotide -7670 of the 3' end of intron 1.

10

15

COMPOSITIONS AND METHODS FOR MODULATING HDL
CHOLESTEROL AND TRIGLYCERIDE LEVELS

Abstract of the Disclosure

- 5 The invention features methods for treating patients having low HDL, a higher than normal triglyceride level, or a cardiovascular disease by administering compounds that modulate ABC1 expression or activity.

50110.004002 utility application.wpd

Figure 1

SEQ ID No: 1

gtctataatggcatgccacagggctctaaaactttgcagttttatcattaactcaaagtgaatgtatacatg
ccgctgactcaacatttttgagagacaacaaatacaatgaatatcaagatacatatatataatgtattt
ctttttgagatggagtttctactgttgtgtccaggctggagtacaatagcacgatcttggctcactgcaacc
tctgcctcccaggttcaagcaattgtcctgcctcagcctcccaagtagctaggattacaggcatgtgccacc
acacctggctaattttgtattttttaagtagagatggggtttcaccatgttggctcaggctggtctcgaactc
ctgacctcaggtgatccacctgcctcagcatctcaaagtgtgggattacagggtgtgagccacccccaccgg
ccatatatatatatttttgagatggagtcttactctgtcaccaggctggagtgaatggcttgatctcggc
tcattgcaacctctgcctcccagttcagatgatttctcctgcctcagcctctcaagaagctgggattacagg
tgcattgccaccatgcccactaatttttatatttcatcatggggtttcaccatgttggccagggtggtgtcg
aactcctgacctcaagtgtctgcctgccttcggcctcccaaagtgtgggattaccggcatgaaccaatac
gcttggcaatatattttttaagaaaaaaaatttcaggttgcaacagcatccaaaaagtaaccaatgatttta
ggtagaagggtgaagacaaatgtaactctttttttttttttgaaatggcgtcttgcctgtcgcctaggct
ggagtgcattggtgcaattgagactcactgcaacctccacctcctggactcaaacgatttctcctgcctcagc
ctcccagtagctgggattataggctcgcgcgcgcacgccccgctaatttttgggttttagtagagacggg
gtttcaccatgttggccaggctggtctcaaactcttgacttcaagtgtccgctgccttggcctcccaaag
tgctgggattacaggcttgagccacccggtagaatgtaaatgttaaaccctgtgttttgaaaatgcaatag
tataggataaggagaaatgactttctgaagaccagaacatttttagtcaatttcaaacacaaatgtgagtaa
ttgtataaaacaggttcttctcctgatgaggataagaatagtatccttgtcagatggaaatgccattcag
ctgtactttctagtggttacgcccatagtagcactgttgatggaaccaggtatctgactttaggaaagatgt
tccccaaactggagctgacccagaggagcctgaccaacttggggaaagttaaagatctcatcacgtggagaa
taggggaaggcaccaacagctatttgagtgtctactttgagcttaaggagagaaggagaaaaggcagggaataa
acggaggtggaataagaataggtaattcttcttaggttaaataataagtgttgcctataggaaggagccc
cagaacacagttatcaataatagagactcacacagagcattctacactagagctgtgtcctcttgaccaga
ataagggttaagggtgtgtgtgcgtgtccaggaaagttaggcagctaggaggtgatcagagcatacactactgcc
gccacaattctaagtgtcttcccctaggggaatcctatttcttctcaggcacatttgtttattcattccatg
ttcactcttgttatttacttcttgcaggctttgtgttaagaattggggaaacaaggtgataaaaccaggt
ctgtaaagaaaaggagctcacagtctggagggggcaaatgggcattgtgcctgcaagttggcccactgagagc
ctaagaagtgaagttatgaatccaggattactcagttatcaatgaagtgattaaacatcatccatacagacc
ttcagagctggaggggaatttggatacctactcagcacatagttttcaaacagtgccctgtggaaccctagg
gcatttcttagggattgctgtgtgtgagagaggagattgaatcagaagggtgtctgggaccattcttactca
catttcaagcagagcagctccacttctatctgtattattttttattttttatttttttztatttttttga
gacggagtctcgcctgtgcgccaggctggagtgcagtgccagcatctcagctcactgcaacctctgcctcc
cgggttcaagagattctcctgcctcagcctcctgagtagctggaattataggcctatgccaccacgcaccagc
taatttttgtatttttagtagagacggggtttcaccatgttggctcaggctggtctcgaactcctgacctgt
gatccaccggccttgcctcccaaagtgtgggattacaggcattagccaccatgccccgcctatctgtatt
atttattcattatttgcattgtgaatgaacctgaagaatgcttactgttactgctaagtatttaaccacaccc
catgcccattgcaggatgatagtgaatagtgcccaaaagataactataattagactcatgtaattaaaggaatat
ttttgtctgtactattatgtgcctataaaagactatgaaatctattttattcagtgattttattggaatacca
aataagcaaatgctctatgtgtctaaagattctaatttgtgctaagattttccttcagatgtttggctttct
caaatccctgagggctagaactttgcctactcatttgtgtttcccaagtgtctaacgcagtgccctgacac
atacaggatctccaaacgcttgcctgaatgtgtgaggaaggaattaaaataatgtaccgccccgggcaagtgcc
tcatgcctataatcccagcactttgagagaccgaggtgggcagatcacttgaggtcaggagctcgagcccag
cctggccaacatggtgaaacccgctcttactaagaatacaaaaattagccgcgagttgtggcaggcgccctg
taatcccagctactcggaaggctgaagtgggagaatcacttgaacctgggaagcagaggtcgtagtgagccg
agatcagaccgctgagctccagcctgggtgacagagcaagactccatctcgaaaaataaaagtaaaataaaa
taatgtactaactggaccagaacagattttccaattgattattgacaacaaaggaatctgaattatttaaat
aagggtgaataagtacatattcatatatatatgtatatgtgtgtgtattttacattttttataaaagtgtaaaag
tatatatacttttttcttttcttcaggtagaaacctctcctagattgtcactgaataaacatttagcactaa
ctatggcaatcaaatacacatattgattgggtgtcagagaagaattgaacattcaactctgaagcagtgattt
tcttccatctgcaaacactctgtcccccatccttcttgtgtatcctggaatccaagtcataaaataatgat
aggattttctgtcaaaggattttctcagaggagtgatgttaactcctttcctctgatacactgactcact
aagcagctacccttgtgaattccaattagcaatacttcttgcctatgtctggtccaactttcagacaaacct
gtgttcaggattcctatagccatttatagggtgtagacaggaagcattcagatatccccagagggtacctgata
accagctgatccatgactgtctgtcttgggctggccagcttgaaatcttgacattgtgggttctczcccaga
gaagggtgccttttggatgtgagataaagacattatgactagatagtgcatgggtgaggggtgtttttctagttt

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

taccgaagtgtgatctgtaaagctgctaccagcacacaczzazzcacacacacacacacacacacacacacac
aattaaccacagatatcctcatgggaaattgtcttaggaaagaatggaagccaagattttatatatagacca
cagaaggtgatgggtaattgttcttgggaaggaggttgacaggcaatagctataagttaactcaggaaagcaaa
gaaaatccccaagagctaaggagaggttagagattctgctttttatttagcaattcatagctctcaagttt
catacagtcctttaaggctccctcttcatatagaataaatgaaattttttataaattgttcctcagattcg
tatctgtacattctgggaccacgagttgtagcaggatgtgattttcctcattctgggcatctaagttctaca
gttaaggacactgaaacaaaccttttagtgaataaagattggcacattgtttcttctcccataaacattgaa
tggtccaggaagggccaggtgtggtggctcacacctgtaatcccagcactttgggaggccaaggcaggtaga
tccccgaggtcaggagttcaagaccagcctggccaagatggtgaaacctgtctctactaaaaatacaaaa
aggaggagtggtagtcccagcttttgcagtaattgtccacattgggactttatcttaaatgtgcaagacaaa
ctcttggccaggcgtgggtgaaccgagattgtgcccactgcaactccagccagagaatgcaactccagcctgagca
acagagccagactccttctcagaaaaaaaaaaaaaaaaaaaaagacaactcttaaatgggtgttctactgtc
cttgaaccaccatacctttgcagaagtcattctagatcatcaatctgaccataaaacacagtttgccaaact
ggcaactaactgcacctatttgtgattagtttgggaagaaacttgaaagtggcccttttaactgtccactat
tccctgctgggacggaattctctcccacgcttaataacggacacttttaaaaaattttttactccccac
agagttggcagattgtgtttcagagagttaaatggaatgctccagtggaaagtatcccgttttctagaatg
gaaagtctatcttcacagtgctcataatccaggtgcccctgggctgagacttcccctgcttaggcgggtacccctg
ggtagcacagctgaactggctgtgaactaaacattcattttttatttagcagccaggctggacagagatcaca
caagaccaacctgacaacagcagcatgctgctcctgcttcagaagtaatttttttttttttgagacaagagc
cttgcctctatcaccacaggcaagtgcaagtggcatgatctcggtcactgcaacctccaccttctgggatcaa
cgacttctcctgctcagcctcccaagtagctgggattacaggcgcatgccaccatgcctggctaatttttt
gkacttttgtatttttttttttttttgagatggagtcctcgctctgtcgcccggttgagtggtgagcca
ccgcacctggcccttcagaagtaatttttaaggaaggatcttgcctgggtggggggttgggcaaggctgtg
aaaagcaaacaaaaatcacatgtggcctaataaggaggccagtggaatacacatgatgaaaaagaaacttac
aaaagcacattattaattttctgaacatgctaataccatccaataacaataagatctaattttattgaatgc
tgattactcatcagctacattccaagtacttctacatgtacagttatgtgacacataagatgttttaggtcac
cactgacagaatgtatgatggtgaccccataagattacaaaaattcctattgcctagtgcataatgtagcaa
gatgcattactcacgtgtaataagtaattgtgatgctgctgtaaacaaacctactgtgctgccatgcctatag
aagtctaggacgggctgggacagtggtcctcatgcctgtaatcccagtagggtgggaggccaaggcaggccga
ttgcttgagcccagaagttcaagaccagccggggcaacatagtgagagcctgtctctacaaaaatacaaga
acaaattagccaggcatggtggcgcatgcctttggtcccagctactcaggaggctgaggcaggaggatcact
taagcccaggaggttgagtctgcagtgcagcatgatggtgcccactgcaactccaacctggagagagagtga
cccggttttaaaaaaaaaaaaaagcctaacaaataaaaagtatgtatagtagacaataacttggaatgataaa
cagtatgttactggtttatgtgtttactatacttttttttttagagtatactcctacttttaaaaaaga
aaaaagaaagttaactgtaaacaggccttagacaggtccttccaggaggtattccagaagaaaggcattgtta
tcacaggagatgaaagctccatgcattgactgccttgaagacctccagtgaggacaagatgtgaggtgga
agtctatgatagtgatgactgacctgtgttaggcctaggctaatatgtgtgttttaacatattagtttt
taacaacaaagtttaaatcagttaaaaacaatttttttaatatagaaaaagctcatagaataaggatatzaaga
aazgacaatgttttgtttgtttgtttgttttttgagacagagtcctcgctzccatcacctgggctggag
tgcagtggcacaaatcttggctcagtgcacctctgattcccaggttcaagtaattctcatgcctcaacctcc
caagtagctgggattacaggcgcccatcaccacacccaactaacttttgtcttttttagtagagatggggtt
catcataattggccaggcttgtcttgaactcctgacctcaggtgatccacctcctcggcctccaaagtgt
gggattacgggtgtgaagccatcgtgcccgccagaaagaaaatgtttgtacagtatgtacgatgtatttgtt
ttaagctaagtgttattatgaatgaatcatgtgtggtgtagtggtcatgccgtaatcgcatcactttggg
agaccgaggacggaggatagcttgagtccagagttctagaccagcctgaacaacgtggtgagacctcgtctc
tacgaaaatcatcaaaaattatccatgcgtgggggtacatacctatagtcctagctactcaggagactgagg
tgaggggattgcttgaactcaagaagttgaagctgcagtgcagctataattgtaccactgcatttcggcctag
gtgaccccgccctcaaagaataataaaaagctgggcacgggtggctcacgcctgtaatcccagcactttggg
aggccaaggcgggaggatcacttgaggtcaggagtttgagaccagcctggctgacatggtgagaccccgctg
ctactaaatacaaaaaatttagccggtcgtgggtgtgtgctgtaatcccagctactcaggaaagctgaggt
gggagaattgcttgaacctgggaggcagaggttgagtgagccgagatagcgcactgcactccagcctgag
caaaaaacaataaataaataaagtttgaaaatttaaaatgtttataaaaattaaaaagttacagtgcagta
agatttagtattaaagaatttttttataaacatgtagtgttaactctacagtggttaataaactctacagtagt
gtacagtgcacatcctaggccttcacattcactcatcactcactgactcaccagagcaactccagtcctga
aagctccattcatggtaagtttctatacaggcgtgccatttaaaaaatcttttagggcaggcgccggcgt
cacgzctgtaatcccagcactttgggaggctgaggcaggcggtcacgagctcaggagatcgagaccatcct
ggctaacacgggtgaaaccccgctctctactaaaaatacaaaaagattagctgggtgtggtgggtgggcgctgta

gtcccagctactcgggaggctgaggcaggagaatggcatgaacccaggaggcggagcttgacagtgagctga
gatcacaccactgcgctccagcctgggccacagagcaagactccatctcaaaaaaaaaaaaaatcttttata
ctatattcttaatgtaccttttctatgttttagatacacaaataccactgtattacaattgcctactgtattc
agtacagtaacatgctatatggttacgtagcctaggaacaataggetatattgttcctaggtatagggatgt
ggatccctataccatctaggtttgtgtaagtttattttatgatgtttgcatgacgacagagtcacctaagga
ctcatttcttagaatatattttagttagtaagcaatgcatgactctattgactcatgaattcttaccacagacc
tatggggcagctactattgttaccctcattttataaaatgataaaaactgaggtacagagacagtaaaataacttg
accacgggtcattcagctactcaagagtcagggtgggatttaaaaccagatcacatgggttccagagtgttca
cacttacctactatactgtctcaagagcaaggatgttttggttcacttgacaaatgaagatagggacctctt
tcattataagcctatttttaggctaaaatagaagggaaggggacacagtgaatccaggccttctggcatggct
cctcagccctttctgagctggcctgggacagccttccctacctcactgatgccacttccctactgagcgacttt
cctgcctacctcactgatgccacttccctactgagcgactttcctgggctccagacccagtaagcgactttgc
ctgcacccaccttattttgctctactccctgtgcttttatgcctttacccatctgccttggaagctcttct
aacctttgaatgggtaaaggcataaatgtatgctggagaaatcctcagctcagggccaggcacgctggctca
cgctgtaatcccagcactttgggaggccgaggtgggtagatcacctgaggtcgggagttggagaccagcct
gaccaacatggagaaacccctgtctctactaaaaatacaaaattagctgggcgtgatggcacatgcctgtaa
tcccacagctactcgggaggctgaggcaggagaattgcttgaacctgggaggcggaggttgacgtgagccaa
gattgcgctctgcactccagcctgggacagagctagactctgcctcaaaaaaaaaaaaaaagga
aaaaaaaaagaaaaaaagaaatcctcagctgagttgtcaactcctcttgaaactttctcagaccttccaggc
tgagtcgtcgctcattttgtgcttccctcagttcctgggcttctaccttcttcatagcttgtttcatgtaattgt
aattcttacttgcctttctccttcttaagctgagagctacttcaaagcatgggtaggacctagcacgggtgt
atgggacatgggtgggtaccccgtaaatgtttactgaaaaaaaaatgcctaaagcaattgttaacatcatcag
atagataattatgggcatcagagattctgtcttcaagcttatataaagaacttatttttggctctaattat
cctgataattttctcattactttcacttattgtggcttgtggatcaattgttgacattttataaacatttca
ctatttgacaatgatgatactaaaatacgaattaagcaaccattctaaagatagtgatgatgataacatata
cgctggtaacatctttattttcagccgtatcatggaatcctctgtttccattctgtctaggtaggcaggatg
caggtagaacttgtgagaggatatgatttttgtttccatcttagatatgacaggaacttggaaattttgaca
taaatacgaacatccgggattcttaaaacaatctttaaaaatggaatgccttaaaagctggcgagtggt
cacgcttataatcccagcactttgggaggctgaggcaaatggatcacttgagttcaggagttcaagaccagc
ctggccaacatgggtgaaaccccatctctactacaaatacaaaatattagccgggcgtagtggcaggcgctgt
aatcccagctacttgggagcctgaggcaggggaattgcttgaacccaggaggccttgagattgcagtgagc
tgagactgcgccattgcactccagcctgggcaacaagagtgaactccatctccggaaaaaaaaaaaaaa
aaaaggaatgcctttgggaataatttattttataatttatgtataacatatagacaaaccattagtttgcctt
atattttactaaatataaattttagtaaatataaatatttactaaatataaaaaactcttagattttactaaag
agttacaactaattggcctggcgtgggtggtcacacctgtaatccagcacttttaggaggcagaggtgggac
gatcacgaggtcaggagatcaagaccatcctggctaacacgggtgaaactctgtctctactaaaaaaaaaaaa
atacaaaaaattagccgggcgtgggtggcagggccctgtagtcacagctactcaggaggctgaggcaggagaa
tggcgtgaactaagcagaggcttccctaaaagtgatcttcaggataaaaggcagaggaagaggctccatgact
gggttggtgttgaggagagccagagaagcaagctacagaaaagagaaaaaattaatatgcaagagagtaaa
caacacgaagaaaaaacccagtggtggaacactacacgtgagaaagggtgtctgtaaggatgttctacaaa
gcaaatgcttggatattaattcattgcagcaggagatggtaagcctcatgataaagaaggagaaaaaatcaa
gtcaagggtctgaggtactgacccagggtatacttgactatgccagcaactgtttagggggagatttgagct
acacttgtagcaaggcaaaaatctgtaattagttgtaactcttttttttttttgagatgggtgtctcgctctgt
ccccaggctggagtgcagtgggtgtgatcttggctcactgcaagctccgcctcctgggttcaagtgattctc
cagcctcagcctcccaagtaattgggactacaggcatgcaccaccatgccagctaattttttgtacttttat
tacagaccatgttttgccatgttcaccaagctgggtctcaagctcctgacctcaagtgatccgtccgcctcgg
ccztcccaaagcgctgagattataggcctgaaccaccgcgcctggcctaaagagatctaattcttagcaag
tttcaccagggtgctctcctcacccccaccccatccttcccacaaagaattagaacaatgtccctactacc
cctgctgtatctctgactttttacttttaaatctcagcagaatattttactaaatgttttgatgtgggtatat
aaaatcatccctgctgacaaggaaacactttttgaaaaaagttttcattatcaaacagtaagtacagctgac
tgccgtgacctttaaccatttctgagctcctccctcattggacttgggtggaggggactgggtaccaataaag
tcaaagtcttaataatttatgcaagtgcctgaagaaatttgaagtgaatatttctatcatcttgaaatgga
gaaagaatctgtaaacagcaaagccagacgcctaaaggaaaaagatttacagattaaaaataagattgcaatc
tggtaaaaatatttgcacacatgtaacagtcagaaagtgaacacttgggttaacawgagcttttawcag
ataaaaaaggaagaataaacattggattttaaacaccgataaacatgaaaagatgtttaatctttatatttt
atttaatccatattattttttagtttaatacaagaaaatataaatcaaaacaataatacattttatatatat
atatatatatatatatatatatatatatatatagtaggcataagttaaagactgataagactgttgaaaagg

gatgaaaaactaggcttactcataccaatatatatctattaggatggctaaagtaaaaaataactgaaaatat
caagtgtctaaaaggatatggagcaattggaaccctcagacatcgctgrtzgagaaaaacaaaatggtacagc
caccctgggagaacagtttagctgtttcttgtaaaagttaaacatgcgcttaccatagactcagcaatctcac
tcttgggtatttatgctaggaaaaggaaaatttatacttgcacacaaaaaacttgtaagtgaatctttatag
cagctctattcataactgccaaaaactgagagaaaatgtcctttaatgtgtgaatggataaaccaactgtgc
aacatccatgtaatgaaatactacttagcaataataataatattaaaaaccccagaaccattgatgcat
gcaacaaatatggataaatctcaaaagcattatgctgagtaaaagaagtcagctcgaaggatttcatactct
aggattccatttatataacattattgaaacgacaaaattatggggacagagaatagatcagcgggtgccagg
ggtttaggtgtgtggagagggtgtggctataaagaacatgcaagggaaatTTTTTTgggagatgaaatggatc
tgtatcctaattatgggtcatggtaacacaaatctatacatgtgttagattcatagaactgtataccaaaag
aaaaaaagtcatttttactctcttaaaatgaaaaaagaaaaagcctgggcattctaacacctgtttgtgag
agtacacattgataccaagttttatgggtgggcaattttgtctataatacggaaagtttgtctgttctattat
tcagcaatcccagttttgcaaaactatgctaagaatctttgggggcccgggcacgggtggttcacgcctgtaa
tctcagcactttgggaggccgaggtggacgatcacctgaggtcaggagtttgagaccagcctggccaacatt
gggcaaccctgtctctactgaaaaatacaaaaatttagccgggcaggtggcgcatgcctgtagtccgagctac
tcgggagggtgaggcaggagaatcacttgaacctgggaggcagaggttgtagtgaaactgagatcgtgccacc
gcactccagcctgggcaacagagtgagaatccgtctcaaaaaaaaaaaaaaaaaaaaaacaaaaacaaaaca
aaaaactttgtgtacgtgtgcaaagagaatacaaagatgatcatggctgcatttttttaaatgactataaaaa
agaggtacaaccagccaggtaaaagtgggtgtgcacctgtagtcccagctactcgggaggggtgaggtgagagga
acacttgagtcaggagtttcaggccagcctgggcaacatagtgagaccctgtccaaaaacaaaaacaaaa
caccaaaatgtctatctgtaggaatttgttttcaagttgtgatacataggtacagtgaaatattatacatctc
atttaaaatgatgataaaatctgtatttgtttacatgaaaaactgtccactataatgtagtgaataatag
attacaacaaatatatatggaataaaacttgtttagaacaatttctagaagaaggtacaatggacagaatta
tctctgggaagtgggtttataatgattctcattttcttctttgtatctttttcatagtctttctacttttgt
tatgtctggacatttgattatgagcatgtattactgatctattttaaaaattgatttttaattttacaaaaa
ctcatgtaaaagcttgaaggttcgcatttttagaccatgttaaaattttctctggatcaaaacagacttattcaa
atatcttgtacctgtcctccaaaattgcctgccaataactacactacaaaagagagcatttagctgcataattt
tttggactgctgagatcaacaatattatttaccatggcttaatttttaccctccagtatgtgtgggttaca
aacactcttccacatttttgaggcattgtctttgatattttaaatgtaaattcagctgtgcgcgggtggctaa
cgctataatcccagcactttgggacgctgaggaaaggatcacttgaggtcaggagtttgagaccagtttag
ctaactgtgtgaaaccccgctctcttttaaaactacaaaaattaaccgggcaggtggcaggcacctgtaatc
ccagctactcaggaggtgaggcaggagaatcacttgaacttgggagacagaggttgacgtgagccgagatc
atgccactgcctccagcctggccacagagcgacactccatctcaaaaaaaaaaaaaaaaaaaaaaag
gccaggcgaggtggtcacgcctgtaatcccagcactttkggagcgccaaggtgssggatcaactgaggtt
gggagttcacgaccagcctgaccaacatgcagaaacccygtctctactaaaaatacaaaattagccgggtgt
ggtggtacatgcctgtaatcccagctactcgggaggtgaggcaggagaattgcttgaaccaggaggtgaa
ggttgtgttgagctgagatcccgcatttgcactccagcctgggcaacgagcaaaactctgtctcaaaaaac
cgaaaaaattcccccaaaaaacaaaaaaacagcaacaacaaaaaatacaataatgtacctgttttagc
ataaagcataattatagcatatggtgattgggaggatgaaatggaaagggtattttactgacttcagaa
attatgtcctgatagattgatggttgatttaataataacttctgtcaagcatctgtttttagaatcaaatt
actatgactctgcagtttcttgaaatctcatagtatcacatctctgtttgcctttgcatgggtttaagaaa
atgaggagtgtaaaacttcaaacttcgttttcattgtattacatttttgatgacacactgggtcatttccta
gaaatataagggtgacaaatattttcacagaaacataagggtgctattatctcattcaatcttaggtcactcaaa
actctttctctccacacattgaagattcatttgggaatgcttttgtcttattgtgcacccccagtgagggt
gtggttaagtgtttttcattttgtcttctttgtttatctacaggggtccattcaataaaacaaagggtggtg
tcaaacttcaggctcttatgggtttggatgtaatctttgggtctcatttttagttaccaacagagagtggtgc
ttctgacctctttgactcttccctgctgaatttactatgcctttgatacttgtgaagggtgagattttcgag
gagtagtgtgtttttgttagaggttgtaatgtctttcttgcctttgtgattcaagttgtgttcagttacaa
tcataagcatgtgcctaaaaaaatcagatgcaaaactagcaaaagtagaaaactcagggtgacagtcctttaaga
aaagatgcaattcttggggctgggtgcgggtggttatgcctataatcccagcatgttgggaggccaagggtgg
gcagatcgcccgaggtcaggagttcgagaccagcctggccaatgtggtgaaacccgtctctactaaaaata
caaaaattagctggcgctgctgggtgggtgctgtaatccagctactcaggaggtgaggcaggagaattgc
tggaaactgggaggtggaggttgagtgagcctagattgcgccattgcactccagcctgggcaacaagagcg
aaactctgtctcaaaaaaaagagatgcaactcttattactgacacagaaatgaaaatttag
ttacatagttatgtaaaaggactatcagctaggttttagccttaccagatttaggttaattcatttcctgcta
cactcatattctcagccacttccctcatcacattttcaggggtgcagtatataatagcgtcaactcgtgtaat
tccccctactccccatgaacttctaggccaaggggccacacgggggtggggcatatagtataaaggagtaagg

cagactgttggagaaaaacagggattagggcagggtgcgggtgggtcacgtgtgtaatcccagcacttttaggag
gctgagggcgggcgatcacgaggtcaagagtttgagaccagggtgaccaacatgggtgaaaccctgtctgtat
taaaaaatacaaaaaattagctgggcacgggtggcaggcgctgtaatcccagctactcaggagggtgagggcagg
agaattgcttgaacctaggaggcggcggttgacgtgagccaagatcacaccactgcactccagcctgggtga
cagagtgaactctgtgtcaaaaaacaaaaataatacaaaaaacaaaaacaaaaacaaaaacaaaaaac
aaggattagttaggactttgcaaagtgcacacagcatgctccatctaaagagagcttttcagtactagctga
ttgttgttgtgtggaaatacgtcttgtatgggtgagatcttccaccttttcaacagaaatgagaaacacaaac
atctctgtgtgacacttccaaacttgttggctatttctgtttttctttttttctttctttctttctttttt
tttttgaggcagagttttgctctcgttggccagggtggagtgcacatgacgtgatcttgggtcgccgcaacct
ctgcctcctggattcaagcattctcctgcctcagcctccaagtagctgagacatgtgcccgcctgacctga
ctcattttgtatttttagtagagacggggtttctccatgttgggtcagggtggtctcgaactcctgacctcag
gtgatccatctgcctaggcctctcaaagtgcagggtattatgggtgtgagccactgcgccggcgccctagtgg
ttatttcaacaaaaataactaacactcttcagcccaaacatgtgtcagagagctgggtgtggccagtgggccc
cctgggtgtgcaactagtcttcagtgtatgctactgtatgctagtgcagcacaactgaccactgaaaaatagg
tgcacgatggtgaacagattggatttcccgtatctctagaaccccagggtttgagggcaattctgatagat
ggagggggagcgagazgagagaaagaggaggaggaggtatcacatgaaagagatactggactttctgtaaat
aaatgggtgggtatttgggtgagtagataggaattaaaaagaaatagagagatgggcgtgggtggctgtg
cctgtaatcccagtgacttgggagggtgaggtgggagtattgtttgagaccaggagtttgagatcagctctgg
gcagcatagtgcagccccgtccctaaacataaaaattttaaaaaatcagzctaagcgtgggtgggtgcgtgcctg
taatcccagctacttaggagggtgaggcaggaggtcccttgagcccaggagttagazgggtgcagtgagct
atgattgtgcccattttacactagcctgggtgagaaagcaagactcccacatctgagacagagagagaggaga
gaggcagaccatatttgaactccagggtgggtgaataattcatactgattaatttaattgaaaaatatttattt
aaaatttgcagatcggagttttgaaagggttacatttattaaattattgttggactatgaggaacacattgt
tgcttggaaataatatttaattatgatgatgatgatgataataaatgacaataataacaaactatcattt
attgagcacataccatataaccaggcccagtagccaagcagttcacatatattatacatctaatacccagagca
actctataaggcagctctatatgttgccttcatttgaaagatgaataaactgaggtacagagatgtgaagta
acatgytgatggacacccagatatkatytkcagataagaaggataagtcagaagatcagataagcagaaagg
tttmmacacacaggactytcatttagtagtaactyttatgattgggtctaaaattctattttcttttacagt
ttacctactcatcaaatctgtcttaactatttttgcctctccaaattctcttctgttctttkkctactatgta
cacatgcttttaaaatatgttaggatttgcgtgttccatgtgktaatttcttgtttgcctttcttcatttat
aaactctttctctactaattatttgacaaactctatttttcttttttttttgaaaaactctatttttcat
atcatttttctaaatttaaaaaatagcatataagggtgggcaggtgggtcacacctgtaatcccagcacttt
gggagggtgagggcgggtggatcacctgaggttgggagtttgagaccagcctggccaacatggtgaaaccctg
tctcaactaaaattacaaaaaaaaaaggttagccgggcaggtgatgggcacctgtagtcccagctactcag
gagattgagggcaggagaatcgcttgaacccgggagggtggaggttgacgtgagccgagatggggcactgcat
tccaacctgggcaacagagtgagacctgtctcaaaaaataaagtatatatatatatatatataataatt
tacgt
gtaaaaacacagacatatatgaaaaaattacctataatcttaccatccacagggtgattgctgttagcacctat
ctgggattttttttttctttttttcttt
atgaaaaagggttatcaatatagtaacagttattttttattgaaatgttaaccatacacataaataatagctttg
cttttttttttcaattcgcatttaccatgagcagtttctctatatcacaaactctttgaaaacattattttcac
ctttcatggctatataaacattctatcatgtgaatttatcaatgtgtatttttagaattgttatcasatgatta
agcatttttaaaatgttccactcatcaataaaatagtgcacacattcttgaacataagttttttttaccatac
tatcggatatttcttagtataaattcctccaagtaaaattagtgggccaaaggcttttttttttttgaggcg
gagtttcgctctcgttggccagggtggagtgcacatgggtgtgatctcggttcaccacaacctcctcctcccag
gttcaagcaattctcctgcctcagcctcctgagtagctgggattacaggcatgcactaccatgcctggctaa
ttttgtagttttagtagagacagggtttctccatgttgggtcagggtggtctcaaacctcctgacctcagggtga
tctgcacgtcctcggcctccaaagtgcctgggattacaggcatgagccactgggtcctggaccaaaagcctcttg
atacattttgctaaactatcagaatgttgggcctctttgcacctactaacacctcagtttttatgagat
gattacattattttccctctttttcccaacctcctattctattttctttagataccttaagactgcaccccc
aacttagctttctgttctcattcaggcagtggtgtccagggttccaatcccagtggtgagtaggagcagtg
ggggttctctgtgaactctgttaggtaattgagccactcaagggggaaaccagctttgggggtccgctctgtg
gcaagcccaggactctatzgaatggagctcattcctcactaattttacacaccaggacaagtagccazgctaa
gaaattgcctcaatttctactggtcgtcagtcacaaaaagaatgaagatacagaaggaaaaataaaaggcctgc
tcccacagtggggtcatagcacatctccacatttaacctttgacaaggcaaaagtatttctctgtccagac
aattccagaagtcataaggaatactgggcaaggatgggtcaaaacacacagactattgaggggcaacaaagg
aaaaacatgcactcctgaagaaatagatcccaaatgctgaacgagaaagaggattgtctctctagaagtc

taatagaaaacagtatctggaacattaaacgtaggatttttttttttctaatgacttactctcttattaat
atgtcagagaaaaagaatagctcctggctaagaaatacaacagtcctcatccagaaatcacagccaggaat
atggattcttaagtgtcaaaaaagtactctgaataaggaaagaaacgcagatgcactacttctattataata
agcgcttgatttcttttaacccctgagttccaattattttctggcaaatttaagggtagtactgactgcttctg
tgaatctattgttacacttgataatggatctgagttgggggtaatatgtgcctatcaatttggatacttaaa
aatctctctctcttttctcatataccctctatctcacaactttccatttaatgaggggaagtgaatttctttt
ttctgccccctcttttctactgcttctagaataaaaagcataacaggacgacaggagtggagatgagaggaagc
atltccaagcaatgggaaagtatgatgagagtcgtgagttggtagaatgggggtgagtaaggggtgggggtgt
ggatggatggaagaggatggaggaggaagcaggtcatatgatcaggcttcaaaggccttccaatatcttgtt
ttcagaggctttgtaggtcttcttacagaactttgaggggtatttttcccttgtgagcaaaggagaacctag
taaattttgagggatgggctgggcgcgggtggctcacgcctgtaatcccagcactgtggcaggcagaggcagg
tggatcatttgaggtcaggagttcgagaccagcctggccaaaatggtgaaaccccgctctctactgaaaaaac
aaaaagttagttggatgtggttgcattgtgcctataactccagctactcgggcagctgagggcagcagacttgc
tggaaacccgggaggcagatgttgcaatgagttgagattgtgccactgcactccagctctgggagacagagtgga
gactccatctcaaaaaaaaagaaatttttttttttggagggaagaaatatatgattagattgtttttgtgtt
ggtttgccttgtttgttttttttcccaagcaaaaaatcactttactgcaatagggaagacaaataggaagggg
aagaaactagagacaggaacagcagtttggaggttctgcaatacagaagcccagatgtttggcttggactta
gacactgggaatgaaaaataaatgatgaattaaaaataaaaaatatttgggaggtatacttgacctgacctt
gggtgcttttcaaatgagaagagaaagatagagacagatgaaaaagttaaagcaaggatgactatgatttccc
atagtgaatgtctggatgatgatcattaaatgaaatgtttttaaagggcagatggtagaaggagatttga
aggaaagacagaaatatttgttttggatttacttgtagaatgacctgcaaaatttgtcctaataaatg
ttaacaagtggctttccatacaaaaaaccaaaccacaaacccctgatgtaaacaaaaataaatcttgag
gcccccttcaccatctgaatgaacttctcctctgcaagggcactcttaaaatttaacatgaaagactgggt
tcaggatcatgacgggaagtgggggtcgacaggcctcattatgcctctctggcattaacatcaacacagacc
ttaagtctgttaagaagcatttacaatctattctctctgaagcctgctacctgaaggcttctctgcacact
gagaactttggctccacaatcctttatcttaagccagacatttcttttctattgatcccagggtcttttagat
aaactcaaccaattgtcaaccagaaaaattttaaatctatctataacctaagaagccccacttcaagttgcc
ctgcctttttgaactgaaccaatgtatttcttaacttatttgattgaagtctcatatctccctaaaaacca
agctgcacccccaccaccttgggtgcatgttcttaggatctcctgagggtgtctcctgagggtccaagatca
ctcatatttgggtcaccataaatctctaaatattttacagagttttactcttttcatcgacactgatttata
ttggatttccaatggtgtaatatctcatgcccgttggccgatttcaagctagctggttagagatatgcaacaaa
caaacaaacaaaaagagatatgcaaaaaacactcttgtgagcttgtgtgctggctccaacacaccatgtctgc
tggatattctaccaggcagcataggggcgcagggtgggtgagcctggcaaggagacaagagctggaggtacag
ggcttaaatgaagccaaggcgtgaataatattgatgaggagagcaagtagaaatgaaaagagaaaaaccagt
attaggagatctataaagaacaatagtcagaaaaagagtgggttaaagaggcagaagaatcagaagagaacc
agattcatgaaataaaagagaggaatcagtttcaagagaggtcaacagcwcaccaattaagcgggttttctgagcttaagatatt
tttctttatttttctatgttaaaaatattttataaagaaaaagaaaaaagagatcacatttgcaccccccaa
cccgctccctggcctctctctctctaaggaagtgcgttaatgaatttgggcagaatagattcagagctgtgaga
gtgaaaaatttgactgcagtgagggtgaggtgttaataaaagggaggggtagagatgaaatctttccagcagtt
tggctcttaaggggatagcttgaaggaaagaggggttgaagggtgggtttttaaagaattatcactcacta
atcaactagaaatccagtggaatatgcagttacttgtctgtaatccagcaggctaaactttttttttttttt
ttttgagacggagtctcgctctgtcaccaggtggagtgcagtgccatgatcttggcttaccgcaacctct
gcctcccagttcaagcaattctcttgcctcagcctcccaagtagctggaactacaggcatgcaccaccatgc
atggctaaatttttttgaatttttagtagagatgagtttcccatgttgggtcagggtggtctcaaatctct
gacctcaagtgatcaccgcctcaacctcccaagtgctgggattacaggccagagccacagtgcccagcca
aggctaacctcttgatcccaatgacaaacagaaacaaacatcttactcaagtcagaaagcaataaattttg
aatcttgccttgcatgtcaacaggagccacattaatacagaagaggatcacattgggtccaattaaattgaatt
gattgagagcctctgcaatacacggtctactgcacaaataatgatgggtcctgggtacatttttatttgacca
ttgattgctcgatttgtttctgtgtctaactgtgtattggaattaagctgactcaatttgaactgcaggtcc
ttttatccctcttatttttatttatttatttattttaaaggcattttctactctccagaaaaatctct
aagatttcagctattaggcatttgccttt
ttgcccaggctggaatgcagtgctcacaatctcggtcatcacacctccacctccctgattcaagtgattct
cctgcctcagactcccagtagctgtgactatagggtcaaacaccacacagccagctaaatttttgtattttt
agtagagacaggttttctactatgttggccaagctgggtctcaaacctctcaactcaagtgatctgccegcctt
ggcctcccaagtgctgggattataggcatgagccaccgtgccccaccaggtcttcttaatacaacaaatt
catcttaaacatcatttttaaaatatatttttttttattcttcattgacttctgcattactctattttttt

agagtttctagcaaccagtatcattggtatttttaacaatgtgtacatgtacatttatgcagatgagttaac
atatatcaaagcaacctccaaacaatgccatttaggtaatctccaatttaaaagcctcaatagaatgataag
attgagcttttctgtagttccatgacctccagcagagtctgcaaggccacagctgacctgaagggtgattctg
taattagaagatgccagggtcatctcagaatagaacctcaagccaccaggctacatttacagaatcagcct
ctccagaaaaacagcaacaaaggagggccttctatgtatttggaggagtcccttagaggagggacttggg
gttttgggtgttgggtgtggggggcagggatgggatggggagggggaagcttattgaaatatactaaaagaca
accaacctaaagggtggagggaagaaaattcacacttgtaagcttcttttttaaggggcatctcttaggctc
tagcttttgagattcagtatatatatatttttgagtcttgctctgctggagtgcagtgggtgtgatctgggct
cactgcaagttctgctctcaggttcacaccattcttctgctcagcctcccagcagctgggactacaggc
gctgcccacatgccagctaattttttgtatttttagtagagacggggtttcacctgttagccaggatgg
tctcgatctctgatctcatgatccgccgcctcggcctccaaaatgctgggattacaggcgtgagccacc
gcgcccggcccagtatttttgttttatgaagatattacatttgtaaagtatgagcttgggtgtcagcaaaactt
atatccctgtgtacaaactggccgaatcacttagccactttgggccaatcacttagctcttctaacagtaa
gaaatcaacaagaaaaataaacatttcaaacattacaatgtgttcatgtattcactgtgggggatgaccaga
ttctcgaaaccacaggttgttcttagtgaaacaagtttgggttggggccatagacttgtgtatttagaatca
atggctgttctctctctgggactttgatttttttcttgggctcatccctttttgtagtatcttatctttgt
cttatttgtataggacttaactgttcccattcccttatttagagcaatctaagtgtattaccttcataccttt
ggaattatatgtctcaaaattccaaaaagaatgattttgggtgggcacagtgggtcacacctataattcca
gcactttgagaggctgaggtggatcgctgaggtcaggatgttgasrccawcctggccaacgatagtgaac
ccgctctctactaaaaataataaaaaattagccgggcatggtggcaggtgcctgtaatctcagctactcggg
aggtggaggttgcaatgagccagatcgaccattgcactccagcctgggcaacaaaagggtgaaactccatc
tcaattaaaaaaaaataatgattttgggtgtcagcttcaaataggtaggagaagaaggagagaggatgg
agggtcasggagatctaattactctctaaaatcatgctaggaaagataaacaccttttaataaacactctctgc
ttttataacatcattctgccaaggagctcaaaggttcaacamagttcactttcagaaaaccctttgagga
agacagaatatacatcttctctcmttttaagatgaagaaacaggccgggcacaatggctaattgcctgtaa
tcccagcactttkggaggctgaggccasargatcgcttgagctccaragtttgagaccagcctggataacat
ggcaaaaccctgtctctacaaaaaaaatacgaataattagatgggtgtggtggcatgcacctgtggtcccagc
tacttgggaggctaaggtgggaggatcgcttgagcccaggagtcaggtctacactgagccatgattggatc
actgcactccagcctgggttagacagagcaagaccctgtctcaacaaaaatgaatgaaagagaaaagaaa
gagtgagaggagagaggatgaggggaggggagggtagcagggagggggggaggaaggaaggaaggaag
gaaaaaagatgaaaaaagaaatacgcaacatgaaacagaggcagaaagactttacgtaaattgctcatcat
gtggttgtcaagtttgacccccaaaacccaatttattgaccaagggtattctttgactgaggcaagggggtcc
gctctcctgggcttgggctttagaaagctcatctctggcctttctgagatccatccctttctttttatttt
tcttgacacggagtcttgcctgtcactcaggctggagtgagtgccatgatctcgactcactgtaacctct
gcctcccgggttcaagcgattctcctgcctcagcctcctgagataaacaggcgctcgccaccacatctggcta
atttttgtatttttagtaaaagactgggtttcatcatgttggccaggttgggtttgaaactcctgactgaggt
gagctgcctcacttggcctcccaagtgctgggattacaggcatgagccactgcgccagctcagatccatc
cctttctaagggcaaacagttccatggtgcaaaggggcatgccaccagagttatgagtacctgggactcca
gaattccttgctggtggcctccacatgcacttccagggcctgcttgggctcttctatgggtctgtcctga
gtgttgatagaaccactgatgtgagtacctgggcttgagcctggcctggagatcctgttgactgtagcatg
gagggggttgtgcagctgaatgtctgyatgcaggtggtgggagtcttggaaatgatggagctggaggtgg
gaagagaagtaggcttggggcagctctctcatgccacctcattctggccaaaactcaggtcaaactgtgaag
agtctaaatgtgaatctgcccttcaaggtggctacaaaggtatctttgtcaaggtaggagaccttgtggcct
ccagctgcacttccagggcctgcttgggctcttctacgggtctgtcctgagctcttctatgzaatctgtcct
tcagggcagattcatatttagactcttcacagtttgacctgagttttggccagaataagggtgacatttagtt
tgttggcttgatggatgacttaaatatttagacatatggtgtgtaggcctgcattcctactcttgccttttt
ttttgcccctccagtgttttgggtagttttgctcccctacagccaaaggcaaacagakaagttggaggtctg
gagtggtacataattttacacgactgcaattctctggctgcacttcacaaatgtatacaaaactaaatacaa
gtcctgtgtttttatcacagggaggtgatcaatataatgaaattaaaagggggctgggtccatattgttctg
tgtttttgtttgtttgttttzztzztzztztgtttttgtggcctccttctctcaatttatgaagag
aagcagtaagatgttctctcgggtcctctgagggacctggggagctcaggtcgggaatctccaaggcagta
ggtcgctatcaaaaatcaaagtcagggttgtgggggggaaacaaaagcagccattaccagaggactgt
ccgccttcccctcaccagcctaggcctttgaaagggaacaaaagacaaagataatgattggcgtcctga
gggagattcagcctagagctctctccccaatccctccctccggctgaggaaactaacaaggaaaaaaa
aattgcggaagcaggatttagaggaagcaaattccactggtgcccttggctgcggggaacgtggactagag
agtctgcggcgcagccccagcccagcgttcccgcgctcttaggccggcgggccggggggggaagggg
acgcagaccgagaccctaagacacctgctgtaccctccazzzczccccacccccacccactcccccaac

tccctagatgtgtcgtggcggtgaacgtcgccggtttaagggcgggccccgggtccacgtgctttctgc
tgagtgaactaactacataaacagaggccgggaacggggcgaggaggagAGAGCAGAGGCTTTGACCGAT
AGTAACCTGTCGCGTCGGTGCAGCCGAATCTATAAAAGGAAGTAGTCCCGGCAAAAACCCGTAATTGCGAG
CGAGAGTGAGTGGGGCCGGGACCCGCAGAGCCGAGCCGACCCTTCTCTCCCGGGCTGCGGCAGGGCAGGGCG
GGGAGCTCCGCGCACCAACAGAGCCGGTCTCAGGGCGCTTGTCTCCTTGTTTTTTCCCGGTTCTGTTTTT
TCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAGAAAAGAGACGCAMACACAAAAGTGGAAAACAGgtaaga
ggctctccagtgaacttacttgggcgttattgttttgtttcgaggccaaggaggcttcgggaagtgtcgggt
tcggggactttgatccggagccccacatccccaccacttgcaactcagatgggaccggaggcggtgttaaat
ggggagacgatgtcctagtacgagctctggtgaccccaggactctgcgctgctgcgcttgzgggcttgcccg
acggtggagaccggggagcatctctgggcgtggagaccggcgagtagtaccocgggctcagaggggtcgggg
gttcccgzgcgtgctgagggcgctgctgcgggtggggagagctgcaggtccggcaccgagzcgctgcttt
gttcggagggccctgagctggcztagzaaaccttctggttcaggtcgccagtagctacagagacaaatg
ccagcactgagtcctcactcggttcttaagaagctggtctgttctgacctgggaattggctatatgtcccc
gggactggagcggcacagtcctcgactgtgaatccgggaactcgagttggaggtgtccaaacgggtcgtgg
tgctattgtcactagaggccttgggtctttgztgttgacctgaggggtaggaggtcctgctacagtctcc
gtgcgctcagctgagctggtgtccctggcgacagcgcgacgagttttgtttccttttcttttctttttt
ttctttcctttaagtctcggtctgtcgccaggctggagtgcaatggaacgzatctccgctcactgcaacct
ccgctcccggttcaagcgattctcctgctcagcctcctgagtagctgggattacagggcgctcaccaca
tcagctaatttttgtatttttagtagagacggggtttcaactgttgccaggctggtctcgaacctcga
cctcaggtgatccaccgzgcctcgccctcccaagtgtctgggattataggcgtgagccaccgcgcggccg
agttttgtttcttttaaaaacaagacttaggagagcctcgagagaccggaggtggggtgccaatcctccc
tctcccacgttccctgcagccccatcttcagaccgttgcgtgctggtctctcggggcagcttctgcctgggc
gcagatggggaagctgggcccaggtggtgctggaatgaccgggagtaaccccgcgggcggcgcagaact
cggagctccgcgcgggggtggtgctggtctgctgaggggtgggggtgctgggcgcgcgggctgcgggtg
zccccggagactggccgcagzgcctcctggccggaggacctaggaatcgcccggtctactaggtgtctt
tgctcgcggttccgactgtgaatccggtgaagaccggtggttcagacggggaggaaactatgaggttgaggc
gaaagcccgttttgtttttttttttttgtttttttgtttttttttgttagcgtgtttgccaactcccag
gccattggtaaagcaggaaggttcttggggcgccggacgggtgccagggttatgtgtagggtgcctcttaggt
atatcttttatcaaaaagaagcaaaagaaataagattaaaaataaacaagaaaaaagttgtctggcactggc
agtaattggcctgcctttgcagcactgataccattagcttttaaaatccgacttttcattgacacttcaaga
agagaatgggtagtatatacacattcatctcatagtggacaaatttcatatttaaaaaaaccttctgggtac
tgaaatcagcaagtcacttgccctccatggccgaatcctgcttcccacgaagagaacctcacaaaaatctc
ccccagttaaagagtggaaatcttctgatttttttztcttttttttttaacggccgtagtttagaazccc
agacttaaatatgatcttcttttcaaaaaaacttaagtccttaagttttcatctccccttttatttcaa
cctattcttctcatacctaccacaaaaaatggaggctttctggttgagaaactttccgtttctgttgagag
tatcattctcttgagaaactttctcctaaatcagagaaagtatggaagcatggaaagtattcctgagtagaa
cctctacagataattacaataattttcaaatacaaagtttccattgtcagcctgtttcccaagtgttccaca
aaccataataattccacaaaccattaaaataaataatgctagggaaattttaggaaaacattgggtttacaa
tcagaaggacccggggaagtgggtcttcagccttcacgatgactacaagccatttaagggaactagaattgcta
ctgtgtgcagagcaatttaggagctgtgatttgagcaccgcagatgttccagaatgacatatctgactgt
aacctggacacgtgtgatattgtgtctccctgcagatgagcatttgaaatctcaaccctcgtatttctacg
agtgcaggcctataatggaccctgggcacatttttttttttttgagatgcagctctcgctctgtttcccagg
ctggagtgagtgccacgatattggctcactgcaacctccacctcctgggttcaagagattctcctgcctcag
cctcctgagtagctgggactataggcgacgccaccatgcctggctagtttttgcattttcagtagagacag
agtttcaccatgttggccaggatggtctcgatttccctgacctcttgatccaccgcctcggcctcccaaagt
gctgggattacaggcgtgaggcaccgcgcgcgacccctggacacattttgacttagaacatatttccggtt
gtgtgagacagtgcattagtgacagattggaaaagagtgatcaggaattgattgttttcaaggattgggtcc
ttctgctcaaggaagtccattgttaaacataaaaaaatgaatgaaactgaagaagttcagtgacttagcttt
ttattatcttctgtagtacttacctttttggagaggagttggttgggatattttccatttaaattttttt
ttaaagggatcttctctccgtaagccgggatacttaagctatatatgtagtggtacaaatttaaggtcttc
actgttttcattttttagctgctagaataagtgaacattaccttagatagactcttctaatttaaggtcttc
tagatgtctagaaaatatcaaatgcatgtggtttttgtactttctaaaatacttttaaaacaaatactttt
tctttttttttttttctgagatggagctgtgtctttttgcctaggctggagcgcagtagaatgatcttggc
tcaactgcaacttccgcctctcaggttcaggtgattctcctgcctcaacctcctgagtagctgggattacagg
tgctgcccaccacacccggctagttttgtgttttttagtagggacagggtttcaccatgttggccaggctgg
tctcaagctcctgacctcaagtgtctgccagcctcagcctcccaaagtgtgggattacaggcatgagcca
ccacacctggcctcaaatacatttttttaagtatccagatattaaataaataataaccattatagtagttgtt

atgggtcatttactctagcatcaaagtgttaaaagatcattctgaacacttgttttgtttatgctgagagaagg
cctactccaaaaaatgcaaccatcttcgtatctgcatgtggatacaaccatgaatggccaaagttattgcag
tggtgaatagacacttatatagcactgtgtggcaagtactgtttgaaatgtttttcacgtgtttattcattg
tatttattttgagatagggctcttgctctgttgccacaggggtggagtgcagctgcacagacaagggtcactgca
acctcagcctcctgcgtcacgtgatcctcatacatcagcctcttaaggagctggggccacaggccacgcacc
actgctcctggctaaattttttacaattttttgttaaagacaagggtctcactatcttgctcagactgggtcttga
actcctgggctcaagtgatcctcccacatcagcctcccaaagtgtgtggattacaggcataagccactgtgc
agggtcatcaaagtctaattgaattttcacacaaaaaccatttattgtccctagtttacaagattaagtaaat
gagaagctaattttttctctggctatataccttgcaaggaggcagagctaagacttgaacccagccagagtctt
ttaactccagcactaacatttcagctgctgcaaccaggaggcttttcaaggatgatcaccacattctctaca
ttcatctgctataatccttatcagaatctacagcctgtatcatattttccttggtgctgtgagtggtcagc
caaattctctttaacttgaaaccttggttgctgtagggattgcaacatcctggaaagaatagaataaaattta
ctcaactcaattttttacttggttcataatgaaaactatactattgcttcagtcagatgtttgccaatagct
gtgtgatctcaaaatgttttcctatgtgatctatagzaaaatggaatgatagagtattaggctgtaagggcc
taagazaacaaaggaaaaagagaagtgaactgttagtttagttgttaaaccctaactttgggtgaattgtaaaa
atttggtataatacaatatgattccttgcttgctcctgtccttgatgaagttgtggaccttttgaaataagcta
tttctctgttactgctgttactgttttagaaatcaaatttagtttttttctaagatatacgtatttttggag
ataaacacagtttcaaagctctgccttggttggtggtgactggctcattgttgattcccgactttggg
aggccaaagcaggaggatcacttgaggztcaggztagttzcaaagaczagzcctggcaaatatggtgaaaccc
cgtctctactaacaatacaaaaaattagctgggctgggtggtggtgcttgtaatcccgactactgggattgg
gaggtcgaagtgaagaatttgcttgaaacctgggaggcgagggtgcactgagtcgagatcgtgccactttac
tccaaactgggcgacagagtgagactccgtcttgaaaaaaatgtctgccttgtaaaagtgaatataggatga
gaaagtgcctttcttattaattggtgtaattgaattagaataaaactcttgaaagacacctcttggtaaaaat
agttacatttactgttgatttatggtatggttgatgtttttaagttttccgtgtaataactcagttcatt
ctcatgagtgaatatagggtgcttttattgtctttatagatgggaaactgaggtataggcaggctaggtacatt
attatggagttcgtaagtgtggagctgaagtcgcacccagacagtttggtctccgtgagttaccaatct
catggtaaagactttgtcagactatcaaagttttgacaaatgaaatattagcaaaaggccaaaagggttct
ctattttcatttgagtatcttcacctgaaaatagttgcctgaataagtagcctgcatagaaaggtacatttt
agaaatacttgaggccagagaatgaaaagcttacataaaaattgatttccgggtggggccttcagttactctcc
attctacgaagaccacaaatagcattcaggccaaagagcatttattccaacaatggaggagcactggatttg
ttcctaaaaacaaaataaagtttgaaatcctgtctttcccatggttgaaaacaaagttggtacaaaacccctta
gcttttgcaaacctcctttaagacccgatttaaatgcytccctcctcatgaagctcttctggatccactcyt
tcccatcactaagttgaaagtaagatccccctctcttacttccattagacttggttacagcactccttgt
atcatgtattttaattctgttttttaattacagttaacatttatttgtcttctccttgagtgtatcttct
agaggaaggctcttgattcattctccccctggccttaattcactcccacttaatatggaataaaatttaataat
gctgacttgaataagttcaacaaggagaatgggaagctcatgtttgcttctctctctctctctctctctct
agataacagggttaatacagaaaaagcattagaaatagagttatatgagaaacaactgtagtttaaggctagg
ttatggttagactgagaaatttttagtgcatacttaagttatttagggcagggttactttttgtagaacaaacat
ttcagtttcgctcagtttcttttccgtttctgzaggcagctgtgatttaagaaaatgctctagtctgtggca
ttccatattcaagtaacttttgagttgtatattaatttatttttggtaataagagtgcacgactcactaagta
atttagagatttaaacacttttttaaaaaacagtaacttcataatgcattggatctattcttctataaagtct
tttcttgggggggtgtttgttttaaaattccccggtgttttctctgccaatccaacttccaagaagcatttg
aagtcaaaaacattttatctggttagtcttaaaagtcagatattttgtgatagctgggtatttagtttatgata
tttcccaggaagaacttttagtagttgaaccatttatgaaagacttcttgaaagctaccttagagagttga
tttagttcttctctaaataagtaaatagaatatttagtattaggacatcttgaggtatagatgcaaatattggt
gaaaaagaacatggatatcagagtc aaatgaatgtagattggaattctggttattactaatggatatctgac
attaggcaagttgctgatcactcttgctcagtttcatcatctgtaaaatagggtatttgtgtttgtgtata
atgtgaaccgtataatataatgcttgccctatagtgaatttattcatacagagtgttttcagtgatttttaa
agcttgcttttagggcgggcgatggcttctgcctataattccagcactttgggaggccaaaggtgggaggat
catgaggtcaggagttcgagaccagcctgaccagcatggtgaaaccccgctctactaaaaatacaaaaaat
agctgggctggtggtgcaagcctgtaatccagctactcaggaggcttaggcagaagaatcgcttgaaacc
aggaggcagagattgcagtgagccgagatggtgacctgcacagagcgagactccatctcaaaaaacaaaa
caaaaacaaaaaattctgtcttatagtttacttccacatcaaattgtctttatcccatgttacttgcac
tgatattccagacatgaaagaaaaaaagatgataacaatgacagttattaaattaggttccactcttattc
tagatcaccaattcatattactattcagacttggaacattaaatttttagttaaacttttttcaaatatgca
tataattgtcagtggttactatattttggggaagagattgttgacttctttgaagaaagatacggattttct
cttcagaaraaatacacatgggctcatataatccaaattttatgtgtaattacagggtgttcatgaatgcc

acaaatccattaagccatgtgaccttggacaagtcattttacttttctgttttttaggttgttggctctgtaa
aatgatactacttgactttttaagagcccttcaagctcttatgtcctctaaccacaggtctgtattcagaag
aaggggtggtcctttaattagagccatctagagatctgaggaacatgctgggcattagtgtaacataccatg
tggtttttgagaggtaaagaaaaataaccagggaatgcctcagagcattcctgatcagatcgatgacagaa
gaaaggaatgagagggaggagaggaagctgttgaaatttctatttacctgctttgagtgaatgaagatttg
aatcatagaaccagaaggggttctcatctgaaatgcaaaggaaggaggagttggtttaattcaataagtttc
agttgagtaaacatgatttagtgagatactgttcttgcctctgactcaccatttggaaaatctctctaaaat
aaaattggactctccatctcggacatcattttgggtgtaggttttgcttttttttttgagatggagtctcgt
tatgttgcccaggctggagtgcagtggcgcaatctcggctcactgcaacttccgcctcccagattcaagcag
ttctcctgcctcaacctcctgagtagctgggactacaggcgtatgccaczcagtccggctaatttttgtat
tttttaatatagagacgggttttactatgggtggctaggtcgtttaaactcctgacctgtgatctgccc
ccttggtctcccagagtgcgtgggattacagatgtgagccacagtgcggcgcttaagttttacttcttataat
ggactcctgttaagccaataggtgatgaaaggaaaccataaccaactcttcaggctcattcatccttcaaga
atagcatgctagtacctatcctaggaggagaattggactatacctcatgaggatagtttgaagtatctcaga
agacctcactgggggtaggtcggtaagacacaaagctttctaaagcactgtaccaaatgttggtttgaga
gatcataacaaattagaagtggaaagaagaaggagtaaaaggaagaagggtttctggccaggagcagggag
ggggaaggagctgctaggaagatgtttgggtgtcatatccctgttcaccttgccttgcaaaattccttgtag
gatgccagggtgggagtaattgtttttcacaagagtcaaaccacgcttgttttctgaagaagcaagtcttt
ggaggggtgggtggttgaaatctgttggatctggtatttaggtgatacactttgacataaaggzcaacactga
tgcaagcagcagctttccttggaaggcaggggagaaagtgaaggccagactgatgagcttacactgacctg
cagacctctccattccaggcatgtttgggtggcagagttaaccttagttgggggttaggctgttgtctggt
actgtgagagagaaggaagaagaagatatgatattaacaacaataacataatttatatttgaaaattaaat
gcactaatacacctatagtgtctattaacaaaatttatttgggttcagcaaatgtttgttaaacaccaatgtact
gtggaaagtactaggtgctggacagaggtcaaaagactttttaagaatctgccaccattaatgatctctttc
tgcttggcattcaaggctctttgaaataagactgtgaccttcttgatagttttgtcctggattataagaca
catgctcgaaggaaactatagctgggttttctcaccagactgattaacatatagtaggttggtagctgttaa
atgagtctctctctacaggttttctatcttctactttaagaaccttctcctggcattgtttggctcctcattg
ttctggaatctcatgtccatctcatgtatttgccttgggttcacacttgatcttctgctcatattcctcacct
actgaaattttacctcaaccagacctgggtgatggaacacagcatgggctagttcttcttatatatgat
ctcatttaattttcactggaactctgagataggttagcatttcagcccactcaagttgactcaaatttttgta
atcatcaatatattttaaaataaactttatatttgacctgtaatttgaaaaccaatatgagttatcataaatg
aaaggtaatttttaaaaataattaggatgaagacaaattatttttctcacagtctatgtataagataaactat
tggttcccaaaggccccagctgacaatgagacttctcttactttgttgaaaagggaatttagcaagcattaaa
gaggtgtcaaaaagaagactaaacaaaagcttactccttttttttttgagacgagctgtcctgttggccc
aggctggagtgcagtggcaccacctcggctcactgcaacctctcctcctgggttcaagcattctcctgcc
tcagcctccctagtagctgggattacaggtgcactgccaccacccggctaattttttagcagag
atgggggtcgtcatgttggccaggctggtctccaaactcctgacctcaggtgatccaccacctcggcctccc
aaagtgtagaatcacaggtgtgagcgcgccgaccttactccttaagtactaagaatgttatat
ataggctgaagaagtgtgctgaaaagaaccatattttctcatgatgtggttcaatgtttaactgtgcttgtt
catctcctaaaaatcctctgaatatcacttaaatcatcctgtgtaactctccacatttggggaaatactga
gcttgcctattattatattagccccatatttcagatgatgcacctgagccgaggagaagttaataaactgt
tcatggttccatatttggctaattggcagagccagggttcaaaactcttgtctctctgactcccaggtttgtg
cttttcccacttggctgaatttctcatgctacctcctccataaacacctttcctagaacttttaaggaaatgct
tcccctgttctctcatagcatttttaagtgaagttgccaaagtgtcccagtttgacccctaccagcaatgtg
ggagaaaatagcaacataatttttgatgttggggtctagtgttacgggtttcttctgctcttgggatgtatat
tccatggctactgaatacccaagtcccaaacagttttcttagactcagaaggctcatccacttcagcttcttc
atcaaaaagacataattttctggctgggcacgggtggctcacacctgtaatcacagcacttgggaggtggagg
caggagaattgcttgaacccaggaggctgaagttgcagtgcagccagatcgtgccactgcctccagcctgg
gtgacagagcgaaactctgtgtcaagaacaaaaaaaaggacttgttttctgttccattaccacagtggtga
gaatggcgtgctaaatttattctccagctgccattaaactgcaaattaaaaatcttagtctcttgccctcttaa
tccaggcttcttcatactataccagaatttaggataactattacagtgccctttataggagagaagaagaa
attgtgtctgtagatgtctgttctttcagcttaaaatggacactgaaatgttaaatattggactggcctca
tttatttctcctgtctgttgggtccaatttgaatcttaaggcgtctttcaactggaattttttgttctctca
actaaaaattgttcttgaagtttgaatcagaacaaaatcctgaatgttgagggttccctaaaggctgttt
ctttatgcaaaagcctgaaacccgatgttgatgttggctgcttaaaatgaactgtgaatcaaggcaggggtt
ttattttttatttttttttactttaatgattgtgttaattatagtgaaaaccttgagttcacgagaaagaaa
gcctttgggtcaagtatgttttattaagttgtcagcttgttgcaggatttgc aaatttagtggaattagtgc

cattttttcagttttacaattccagtcacattttcacatgatcagagcatggcttttcttctctgtggagcaaata
gagggctgtctgacacttgggtccagtggttccatttaagcagagtggtatgtccctggagctctgcagaga
agggcatggcactctgaccccagatggcactccgttttgggacattgtccaattctagttcatagcatatgt
gaccaacaccagctctcacctgatgtaaacacttagcgctgtgtgcttgggggattggatttgtgtgaattt
ttcaaaactacagttgacagaaggaggctaccaaaaaatgaaacccaataattccatttttgggaattattcc
cacttttgtttccatttttcccactttgttctttggcacacagaatgtttgatttgtgaaaatcttaataaca
gtagtttttttctataaggaacactcagaatcttgataatattggaataatacagatcccttttgtaggatcct
ctcagacctcatataatagagttcatgtagtcaatattttaagaaaaacaccccttaagtttttgtttttcag
aatcacaagtaagtggattttaacttgtgatcttattcccctttcttctcttaatttagtgaggcagccagc
gagaggggtttgttttgggtatttctaagaaggagtttgcttgtaagtttggaggggcaagacttagactctg
tgtctctgtgcttgccctggaactttgattaaattgtcactaaccgagttagctggccctcgccgggctgca
gaaatagaagtgtcttgcacacatgacatatgactgtctcaagagctggctgggtgaaaggagcttctggaga
aggctgccgatactgtatgaactagaazctggacaagagcctggagattggataactcagtttggcgcaagt
aaagggataaaaagtgttaagggtggcaaaattgtatccaggtgtttataggctccctgagttcctgacttga
gcctatctatgggttttagagttcaaggctctttaccagtgctgacaatcttatactctaggttgaacctccg
gggaagggtgcccttgccttgatggcatgtttaccaggggttctagagcctcaatcacagattctctctagctc
acatgaagttaatgaaaatgaatgtgcttccctacaaattagagaggctttgaggaaaaatcagattaaatg
cactcctgcttgaacttatgtttcttagaacacagctggaaattttgtcacacaaaaccttactttcagtg
catttcttgactgggttgttactgtagtgaatctgctttaactatcttttcttatcgctgaggttttacttc
cattctacatgtgatttgtggagcgctgcgtcattgtgggttcagtgtagtgagagtaggaagatgggtgaga
cacagtagcttgttgcacattgcttaatttatcagggatcactgatgagttagtacactagagaagattgta
ggtagagctgaaaagatggaggaattataaggctcagatttctctcttttttttttttttaagatggagtt
tcactcttgttgtccaggctggaattcaatggcatgatctcggtcactgcaacctctgcctcccgggttca
agttagagacttgatgggtctcacttgatgggttctcctggctcagcctcctgagtagctgagattacaggcacc
ctaccatgccagctacttttttgtatttttagtgagatgggggtttatcatgttgaccaggctgggtctcg
aacttctgacctcaggtgatcacctgcctcagcctcccaaagtgcagggttacaggcatgagccactgcgt
ctzggccaaggctcagatttctaatagagatzttctaattggacatagaggctggaggaatgggatzzgaca
zggazaaactgagtzcaggtgccaaaazzaactttagtgaggggccgggtgcgtgggtcazzcgzzcztgt
aatcccagcactttgggaggtgaggcgggtggatcacagaggtcaggagatcgagaccatcctggctaacac
ggcgaaaccccgctctctactaaaaatacaaaaaatttagccgggctgttggcaggcgggtgtcggtagtccc
agctactcgggaggtgaggcaggagaatggcgtgaagccgggagggcggagcttgagtgagccgagatcgc
gccactgcactccagcctgggagacagagccagactccatctaaataaataaataagataaataaataaata
aaataaaaaactttagggaggtggcagtgctatggaggataggtgcaacctctgtgagaatgtagagaaa
atagtataagttagtggtgaggaccccccaagaggggttttatagtaaaacaatggtcagaagtggcaacag
gataccgtataatgctttcacctctaccaatgcactgggtactggagagcgtccagtttgcctctggaagg
cccttttctgtggacaaagaatacagaaaagagattcctttaataaaccaccactctgtgtcccacccttga
tgaatacttctgtgaaattgccagaattaatcatggtaatatgactactgtacacttactttgttccaggaa
ctggataaatgttttacatacattatcagttctatttttggagaagatacaggggctcagagtccctagggt
ttccagagctgggtgagtagcagagtcaggattcaaaccagcttctctgactctaaaacctctttcttcc
tgctgaaacaatttaactcaagacaacaaaggagtttaaggatttggggagttttctgcatggtagaatagacc
caaaggaaaagaaagaaagacagtgactaagatttgggttctgctgccccaccaaagtctttgagcacttct
ataaatataaatccttcaggttgggagaagggtgaacatctgaagacactgattcttcagagatgtaatcca
aaciaagtgtatctttggtgatattggtcactaaaccatttatcccaaaattctcttggaaaaccgtcctataa
cagcaggaacattatccagccaagttttctgcaataaagggttgctgatagaggcttgctgcttgtgt
ttctzgtzagcztzagggtgtttatgaattcactaatcccttcccttcagatcccttttattctgggtgta
tgattgtgactgzaaaaaaattgatttttttctatgacatagaatgttgaaagggtgatttcttttctaga
ggaaagattcttttttctatgtgctacataccccccgaccagggaaggcaaatagtgggtattgtttgct
gaagtcttctcttgaagggtgcttgggtgtttgcttagtggaatcagcaggggaagagaggctatctctzaa
cattttgttagazgtttcttctgtagttctatagtgatgaaacaaggacttggggtacaggacagatctgct
ttcagaaatcctggctcttgtgaggtttagaagccctgagaccatttagctgggtggcaacgggaatgttgag
gggtgataaataggatctttggtgtccaagtatcagtgacatgatatagatggaggttaaacctttaggatct
ccttatttatttgtttgtttatttttgagacagggctgtgactgtagcccagggttgagtagtggtggcatg
atcataactcactgcagcctcaaactcctgggtcaagcgatcctgctgectcagcctcccaagggtgttagg
attacaggtcatgagccgacacccgggtcagatctcctgtaaaattatattgttgacaacatgaagaatta
tgcttctcaaaagctagttatagatttgtacaatattcatagatttcttgtttcagtttttacaattcata
gcccttattttgaaaattagctatttagcaataattttgtctaggaaattggatgtgtattcaagtgaagaa
ggaagtacagttacctattatcttattgtactaacaatcaagtaagtgtgatgcatttgggtactttaaaaa

ctgcacccaagttacagattattggaattaataaaattcactggatctatatatTTTTTaaacggacagtgtg
atagcagaacctcttatagaatzgatagaattcctctggaatgattggataaacttcatttcattccttgactt
ttaccttggaggatttcttacccttttggcttctcaaatttgactattaaaatgttgccctttaaataagg
aacacagtttcaggggggagtagcagcccatgacccttctgcaaggccccctaactcaaggtagtttccttg
gaactgtggtttatggaatgtttcaggagtgtgaggaggtataatTTAAGGCTGTCCTAGCAAGGATACCTT
taaggatagagggccagtagcatctggaggccagaaaagttazaactgaggcagtcagattagcttczaggg
ctcaattaagctgatgggtcagcctgggagaaattgcaggatgactctcaatatccccccccccacag
cagccacgatctgtctgtctttaaactcatgggtgcagtgaacctgttctttccagggtgtcttggccttcagta
aaccttgtttaggcttgcctgaacgtggctaccgatccaaagacacatgatcagagaggcaattagagaaca
gaccttttccaaagcaagcatgttctgttgggcttagaagtttcatgtcctaataattataggacctgtgca
tctctctggagatgaggcacatgagtcatactgtgattcttctgtcttctgtgtcaacatctcatgaataggca
atcagagctttggcaccaatgtattttcagttcatatctgatgtagttaaatccacctcctgctttgtagtt
tactggcaagctgtttttgatataagacatctagaacactgtaaatatataacatttttatttgtctattat
acctcaattacgaaaaagacatctagaagcaacctcatcaagagagatactgaggccgggcatggtagctca
cacttgcaatcccattacttttgggaggctgaggcaggttagatcacttgaggtcaagagttagaaccagcct
ggccaacatgttgaaacctgtctctattaaaaatacaaaaaagttagctgggctgggtgggtgggcacctgt
aatcccagctactccggaggctgaggcaggagaatcacttgaacctgggaggcagaggttgagtgagctga
gatcacaccactgcactccaacctgggcaccagagttagattacatctaaaaataaaaaataagtaataaaa
aagagagataattgatagctgttgttggaaatttcaacttccactctcacttctgttggtaacttttgggaagttg
ttgaacaaagtgaatacacgcacatacacacacatactctctgttgggtttaaaggtttaatagaata
gtgtgcatataatcactgtttttgaaagaggagaattagttgctatctgtacattttgggtatgtgaactat
ttggatagaactctgagaaatgcattcagaacaacaaacaaaatcataggagaaatagctaagtgggaaggg
gcataaagagttgttgaaaaagttatttcttgagaaaccagctctaagctaggcaagtcacttgctttgg
gggaggcctcagcttctctgtctataagattgcagcaggggttagtgagggaatgagcttcaacattccaag
agattttatctactaatacgcacagtcaaatggagcatgactttgtggaagcctctcctctccaccagagg
ggccaatttctctgtcccagtgagatgttgacacttgtatgatccctgcttggagacttccctctctggaa
cctgcccctggctcagggcatgagggtgactgtcacccttcgataggagcccagcactaaagctcatgtgttg
gcagtggttcttgcggaaggaaaaagaccagccagccatttgttactgcacaagcaaacagcttctggtag
ctgtacagatacatgcactttcttccctcactgtgtttccatagacagatttagtgctgtagaagagtagag
ggcagtcacgggaaggagttcctgtttttcttttggctatgccaaatggggaaaaatcctcctatcttgtct
tttttagtgctcatcctctctcccttttcttcttcttataattctcatctctcctctcctggaaatgtgc
atgtcaagttcaaaagggcacaaatgttttgggtgaggaagaggtgggagaaacagtgccaggtgctaactagg
gtcatcatttcccccttcacagccagcttccctgtgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
gtgtgtgtgtgtatttcttttggcagcatcactgaatctgtctgtctgtctggattccaggtttttggtttag
ggaaaagtaaaagtaattttataatcccagctgtcattttagccaccccztttgtgggtagcataatggtcca
ctctctcagttcattgtcctaagatgcttcatcagaagaataaactccaccccgttactctctgtcccc
ttactctgctttatttttcttctgtcaatcctaccaccaccactgtttgaacaaccactattatttgt
ctgtttcccatccctggtagaataggagcccatgaatgaaggaactttgcttctgttgttaccactgaat
ctctaaggatggaacacacctggcatgtgataggcactcgataaataatttgttgggtcatgggcacctt
gcagagttaaaggctgcagttgtttgtggaatttataagtggtaatgaatatttactactattcctcttcca
aggcgatcacacaataatcaggtttacactatccagttcttaggtcttccaagttatgacttgtgaggtat
gttaattatgataatagaaggcagtttatttgggttcagatttattgatgtgtaatttaccacagtaagactt
cccccttacaaaagtatgatgagttttgacaaatggatacacatgtgtatctaccactgccatgctcctttt
cagctctgtcgtccccctccaccatgaccactggtcaccactgcagtgatttctgtcccccttcatttcacctt
ttccagaatgtcatataaatggaatcatgcagtatgtagtttttgtgtctggcttatttttcttagcatta
ggcttttgggattcatccaggttgtcgcatgtaacagtagcttattccttttatggctgagtaagtgtccc
agttttatttatataatttatttatgaggaggtgtctcactctgtcaccacaggtggagtgcggtagcgcgat
ctcagctcactgcaacctccgcctcccaggttcaagcaattctcctgcctcctgagtagctgggattacagg
caccacccgccacgcccactaatttttatatttttagtagagatggggtttcaccatgttggccaggtga
tctcaaactcttgacctcaggtgatccgcccacctctggctcccaaagtgttaggattacaggcatgagcca
ctgtgccagccccagttttatttattcaccagttgatggctttttcgacaactaattgtttccagtttttg
gctattctgtataaggcttctataaatattcacaataacctaggatgggtagactgggtcatataatagtag
tgataaccttagcagaaactgtcaaacattttccaaagtggctcttccattttacaattccacagtgat
ttggtcccagtgctctccatacacatgtagcacttttaatatTTAATTTAGTGGGTATGTAATGATATCTCA
ttgtgggttttaatttgcatttctctgcagctaataatgatgagtggttctgtctatttgggaaggttttaattta
gcagtcgttgtattctgtagatattaataaactcaaaatatcagtggcatttgcagttaaaatttcccttaa
aaaattggccaaaggtttccagcagtcacttctgccatgcccactgtatgaaacaaggctgaggtgtgga

gattgtcacatttttggaaggagtgatccacttgggtgactgatgagaccagagagcgtacgcctcgggct
tgagggtgaggacgggcgggaagtgcactgcatggccctgctggccttgggaggtgcccagtccttagcta
aagctggcagttatgggaaacagacttagattctattacgtttttcaggatgtcccaggagtcacctgggaa
gctcagcagtcctttgtgactttcaagcatatggtagaagctgctgaacacagagctccctctttggggata
at ttgccccaaatcatttaatcaggcttgagaaatgagttaccacaggtccaggagtgctgccacccttgaat
tctgacaccctat tttctcctatccgtctcttaattaattaagcagacatccccaaagtgttacgacaagcca
ggacccttttgcatactaaggaaaacagggatgaaggaaaacagaaatgggtctctgctctgactcagaaggta
gaaatcctcttttcccagccaagtcttcttagggagcacgtaggaagggtctctgaaccacagtgctcagttgca
ggggaggatatacaggaaaggacattgaagaagtggagacctaagtttgagacctaggcattagccaggctag
cagtgcttgaaaaagtgtcttaggacaagagaactcaccagtgaaagtcctcagtggttaggagagcgtgcagca
tattctgagcctgtatacacatctccagggcattgcttagcaggtggggagtggaagagagtaggctggag
tcacagaaggaggccaggtagaccttgggtgagcactggactctatgttcagggtgctgaggagctggcaaaa
ggttttaagtgcggggagaggcatgttcagatatttgggtctagctgagtaactttgggtgctctgtgacaaat
ggttgggagaccagtgaggtggcagttgcggtcatctaggagcaggatcagagtgccctattgactgggatg
actgtgaagtgggatcctttccagccagtaactggaaatgtgtatgagggcagaagtgagtgactgcattt
gaaacattgagaaatctagtacatagtactgtctcttttatatctttttttttttttttttttttgattttgggt
tgtttggttactaacttgaaaaactgatgtgaaatgtccctttggcttcagttacctgagcagaaggggcc
gggcattgccccaaactctcctcttaggacagaattgtctccagtagtgatcattgtgttctgagttgggggag
caaattgtgcaggaggccagggtcagtgccaagggtgggtgggaggaattggagcaggaagcttgccctaagtgt
gcccagcaaaagccaggtagaactttctactgtggtctatgctacttcttagcaaccttctccatgtgctt
cctggagagtccttggagtcagaacctttttcttgaaaccagacactttacttccaagaaaatgctgtcca
agaaaactcatccttcccttcttctcatgaacgttgtgtagaggtgtgtcttctcttcccttgagcttttcc
actcagggttttaggggaggtgatattctatatttgggtttggctctgggtactgcaacactaggctattaag
at ttcaccttactgctttgccccctctatctttccagaaaccacaaatggatttgctagaaataatggaac
gtcctgtttggacaggatataaccattttctcagctagaggatattgttggaatgaagaaagataaatgggga
gaagggaactcacattgctttggcacttaaatgaagccatgtactgtgttgggaaattatttatattatctc
gttgaaatccacagtagaacacagttgaacaccatacaaggtaagtattgtcatccttattttaccatgagga
aattgatgcttagagagcataaagccttggccaggggcacatagttgggaagccggggctaattcatgcctg
ggctctttctgatagttttccttttttaattgtccctcctcattgttaccttggggatttcaagagattca
tgtagcttctaaatcaacgaactgattcctggagagcagcttctgtatgagaaaaatctagctaattattta
tttcagtgctctctggaatgcaagctctgtcctgagccacttagaaaaaatttgggatgacaagcatgtgtc
tcacaatgctgctctgggttgcagtgctgtgctgccagttgtcatcttgaacaaactgatgcagtgctgggt
ttaactcttctctttttggagtaagaaactttggaggcctgtgtccttctagaagtttgctgagcaaatgg
taaggaaaagaaataggtcctaaggcttgactatttcagagaatttcttgatttattggactgtcaatgaat
gaattggaatacatagtggttaggtgtcttttctctcagacactgcaatttctccaatctcttgactttt
ctagaagttttaatccaagtccttgttgggtggtzagataaaaagggtattgttctactagagactgaccttg
gcatggagatctcatttggactcacagatttctagtctagcgttgggttttgatccataacctcgctactgc
attcttagttccttctgctccttgttctcctcatgcccagtggtccaccctacccttggccctactcctctaga
ggccacagtgattcactgagccatttcataagcacagctaggagagttcatggctaccaagtgccagcaggg
ccgaattttcacctgtgtgtcctcccttccatttttcatctctgccccctcccagctttaactttaatat
aactacttgggactattccagcattaaataagggttaactgctggatgggtggctgggatacacagaatgtag
tatcccttgttcacgagaagaccttcttgcctagcatggcaaacagtcctccaaggaggcacctgtgacac
ccaacggagtagggggggcggtgtgttcaggtgcaggtggaacaaggccagaagtgtgcatatgtgctgacca
tgggagccttgtttgtcggtttcacagttgatgcctgagcctgccatagcagacttgtttctccatgggatg
ctgttttctttccagagacacagcgctagggttgtcctcattacctgagagccagggtgtcggtagcattttc
ttgggtgtttactcacactcatctaaggcacgttgtgtgttttccagattaggaaactgctttattgatgggtgc
tttttttttttttttttgagacagagtcctcgctctgtcgccatgctggagtgtagtggcacaaatcttggctc
actgcacctccgctgccaggttcagcgattctcctgcctcagcctcccaagtagctgggactacagggtgcc
tgccaccatgccagctaattttttagttagagacgggggtttcaccgtattggctaggatgggtctc
gatttcttgacctcgtgatccgcctgcctcggcctcccaagtgctgggattataggcttgagccaccacgc
ctggccgatgggtgctttttatcatttgaaggactcagttgtataaaccactgaaaaatagtagtaaggaag
ttcaggaatagataaagtcactccaggcttgaggcaaaatttacaatatgctgctgactttgtatgtaaggg
gaggcattttcttagaazaagzagaggttaggtctctgggattccagtatgccatttccatcctcagtgtttt
tggccacctgagagaggtctattttcagaaatgcattcttccattccagatgataacatctatagaactaaa
atgattaggacataaacacgtagctcctagcctgctgtcggaaacacctcccagtcctcttcttgggtgaa
cccagaggtgggagctgggtgactcatgatccattgagaagcagtcagtgatgcagagctgtgtgttggagggt
ctcagctgagagggtggattagcagtcctcattgggtgtatggctttgcagcaataactgatggctgtttcc

cctcctgctttatctttcagTTAATGACCAGCCACzGGCGTCCCTGCTGTGAGCTCTGGCCGCTGCCTTCCA
GGGC'TCCCGAGCCACACGCTGGGGGTGCTGGCTGAGGGAACATGGCTTGTGGCCTCAGCTGAGGTTGCTGC
TGTGGAAGAACCCTCAC'TTTCAGAAGAAGACAAACAgtaagcttgggtttttcagcagcggggggttctctca
ttttttcttttggtttttgagttggggattggaggaggaggagggaaggaagctgtgttggttttcacac
agggattgatggaatctggctcttatggacacagracgtgtgtggccggatatggcatgtggcttatcatag
agggcagatttgcagccaggtagaaatagtagctttggtttgtgctactgccagggcatgagttctgatccc
taggacctgggtccgaatcgccccctgagcaccctcttttcttttctgtgcagccctgggagccacctggc
tctccaaaagccccctaatgggccccctgtatttctggaagctgtgggtgaagtgaagtagtggccccactctt
agagatcaatactgggtatcttgggtgtcaatctggattctttccttcaggcctggaggaatataataactga
gacttggttttatttctgcagagggttctaagccattcacttcccagatgggccaataatgctttgagtaatc
tggagatcatctttaatgcgaggtgaatggaactcttcacagagggatgtgagggtgtgagagcagagt
aactzctgaaactcagacgtcagctctttgtctctctatctctgaacacccttcttagagatcccatctc
taggatgcatttctctgtagtttagtttctaagtctcttgttctgttctgcttttatttttttctggat
tctaagccagtatccccacttggctgtcttaatgtagcttaacatgtctgtaataaaaatgatcatctttct
gagattcaaagggtataaaggacttggagagaatttcattcagttttctcaaactagaataatgcttgc
actgtctgtaaaagaacaaaagtgtcaaagcatcttttgttactaaatttctttttattatagtgta
cttaaatatttaggaagtzaaaagtaggtataaacttctttaggctgttattatacaactatatgacccat
acatatttacaaattaagtgcagccaaaattgcaaaatcaataccattcaaattaataacttaaatgtgggtg
aggcagctgttgttcaactgaaaccaaattataagttgcagtaggagtaaatgctatcatgctgatcttttg
agtttggccagctctatatztatcatgtgctaatgtagtaattctccacccatttttctacttgtatgacct
aatttgatggcacctgttccatctctcatgagtttgtctacaattatactggtgccaacacaaatcataaacaca
aatataaacttgggctttgaaatcttgtgccagaacttggctttaaagtaagcatttaaaaaatccatagt
gtttattagactttgttttagatgactgttgaaatgaaaacaaagtgtttaaactctcttagagaacttaaa
tataatccctcagcaatatgtatacagatcttctttgagaaaaactgattgtgttcagcctctcatgttac
aaatggggaacctgaattctgaggtctctagtgcagagaacagggactggaatctgtggaztctatctgttt
taataataattgtaaagtataatagataatattatattaataaaaataaaaagcaaacttagaatgagcttc
catgtgtgaggcactaactgattaggtattttaactagatttattccttttaaggccccgcgatgtactgt
tatttccacatgtttagctggggaacgtgctactcagagagggttaagtaacttgtctgaggtccacaccac
taacaaggagcacaggtagggttcaaactccagataatctgactttggagctggcacttaactcaatgtgcc
taatcgcttttcagtggtgtcattattttgcctattctccatctgagaatattgaagtttctgactccttcc
ttgcctttctccctgcctcccggtggttatccccagggtcttgggtgttccagtcctctatgtccgtccttactc
ttattcctttgctacagtgatccagggtcctgcccccttcttatcctggtagagggggccacttgcctgg
gaaattgtctccgccatgggttatccatgttgtgtgtccattagtgagtagtggaagaatcatatcatgtt
ggcaatgaaaggggggctatggctctggggtagtctagtctgaacctcttattttacggatgagaaagctga
ggtacaaaagcagggaagggtatttcttgaggtcaccagccagcaactgagctgcaaccagagctgagatcc
ccaggactagggcccagcctcattctgtcccatcacagtacttttcttccctcctccaaactatttttatt
ttttatttttttgcagctgcttagcagcttgaagttagaagaaagggcagggaaaagggttttcgtgcttag
ccagggaagggaatcctgcaacaggtatgtgggttgggtcattcaaattgggcccagactccactgggtctgtt
gcttctgtctggttattgcagatgggtttaaagtgttaggattagagagataggcaggtttagccaaaggc
agtttctgactgtgtgagcaggttcttttttaaagaagggaagtggtgcaacaccctgacacaaaggggctt
aagttgttataccactgcctgctaacctgttttcttaactctcttctgatttctaaaggaagtataattt
gctgaatcagaaagaaaagtgatttatttcaggttgcctgatgcttagattgttagagttggaagatctggc
ttgcatctgtacagctgacagaactggggctcaggggggcacaggtgccagagttggtcagtcaggaaag
tagcaccagaaccagctctcctggtggccctacagttgcagaccctttttgctttgctctctgtgtatacta
aagcttctatgtctctgaatctcaagttctgactggtagctactttccaatccacctggcttagatttctag
attatattgttttagacgtcagaacctcttaagggttttggggccacttgttagctcacatagtgagaaccag
ccctgcccattaggtagggaagaagttagcagtcctgatagctgttgcctgcagcgtatggatgttcatt
gcacagttcctgtctcctgagatcctggagtgtatacgttggcctcagagcccagcacagagcctggccct
tgggacatgcttagtaagtatttactgaatgagtggaatgtcttaaggcccattagtttgcaggtcttga
ggaggctcccttgcactaggaagaatagaagcatacataaagcctgtgtgctgcccaggaagactagaa
acgctatgttcagcctggagctgaatggtataccccagagcaaccctgttgaaaggcagtgcttgcctttc
attctgtgtcctgggttgcgtggaactcctgggtccctgcctctcctgtacccccattgtgcagactgagg
ggggaccatcagccagggttagtttccgctgttctgttaggcaaagaataaattgaattgagttgtgaaa
gttgggtgcaaagctcagtttgggtccaaagtaacagtttaacttgtgtgggtggcaggtattcagtaaaaac
agggctggggacaggaagggaagagaacttcagagctttcacgatcctcatctgggttttaggctgatccag
aggccaaggtcccatggaacaaactggacaaagtgaggggtggccacatggcctcttttcttttgcctttat
tattaattttctcaaatagatctgactagtcagtggtgggaaaatagtttaattgtgattttttttttttt

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted January 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

aaactgagtcctcactctattgccaggtgagtgagtggtatgatctcagctcgccgcaacctctgcctc
ccgggatcaagcaattgtcatgcctcagcctcccggttagctgggattatgggcacacagcaccagcctgg
ctaatttttgtatttttagtagagacatggttttagcatgttgccaggtggtcttgaactcctgacctca
agtgatccaccacctcagccttccaatctgctgggattacagggcatgagccactgcacccagccagagtac
cactatttgggcattctttaatgaaaaagaatgaactatccaaaaattaaaactcctcatttatgagctttt
agagaattttacagagtagatggaaactctctgcatcctttccccacttctagtttcacctgacacatttct
tccctgtccttactcctgggcccgcagcagtggtcatgattccaatcccagcttgccaccatctgcctcag
tggcctagggaaaactcctttctccagagctttagttttctctctacggaatgaagaaagttaaaacaaata
gacattttattgtttcatttggataaatactatttaagcatctattacttgtggtatgggttagctgggtatat
agtgggtgaagcagctgggcatgagtagtactgctttcgtagagcttacagttcagtgaggccagcagatgtgaaa
cataatcatcacacaaataaaaaatataactatcaactgtgatgaggattatgaaggaaaaaatccggcaact
atggtactggtggttagatactagcaggtgtgggtagggtatttcattttagattgacaggttgtcacattaaag
ctgagagccctgaagttcaagcaatggttagccaggcaaaagatcagaggcttagagatagggaatccattc
caggcagagagactgggggtgcctgtcccctaggtcagggaacagaagaaagccagtgccactggtggagt
aataagactggcgggggatgagttggttagtagacatgaccagatmattagggczaattctcztctggggza
aggagaattztaatttaataattttattttattttattttattttattttattttattttattttttcaa
gacggagtctagttctgtcgcccaggtgagtgagtgagcaatctcggtcactgcaacttttgcctcc
tggtttcaagcgattctzctgcctcagzctcctagtagctgggattacagacgcccaccaccatgcccag
ctaaztttttgtatttttatagagatgcggtttcaccatatggccaggtggtctgaaactcctgaccttg
tgatcctcctacctcggttcccaaagtgtgggtattacagggcgtgaccacagtgccccctgagaatttaa
ttttattttatgtgcaagaggattccctgaggtagtcaggccacattgtctggtgactcttgggatagaggg
aacttgaatgacaaaaggcccaagaaagcaattgtaatcattacatatacatggaccattttatgctgtttc
ttctttcattttaacattatttagtggtgctgttcacatattttctaaatcatcttctgatttagaataatgat
ttctgatgtgtaggctgtgttttatagttttgaaagtaatactttgatattccattacttztcttgattctca
cagcaattctgzaggtgtatgctgtgcaattttctgtttcacagatgaagagagtattgttaataagttaatg
gcccggcatggtggctcacacctataattccagcactttgggagaccaaggtggcggtgacttgaggcca
ggaatttgagaccagcctggtcaatgtggtgacacccatctctactaaaaatacaaaaaattagccaggcgtg
gtagcacttgctgtaatcccagctattttgggaggggtgaggcaggagaatttgcttgaacctggcaggtgga
agttgcagtgagccaagattgcaccactgtactcctgcctgggtgacagagcgagactctgtctcaaaaaat
aaaaagttgctaagaggaggggtgggtatcttttgggtccaaatctactgtgggtgatgaccttgacattcc
tgatagctgtgcagtaatccatttaacacagtttttataagttcaaaazccctgttgccaacatttagattgtt
ccatgtgtgctgttacaataaattactataaagattctatacatttaattcttttattttttgtattatt
tctgtaggccaaaatctgaggaacaggattactaggttgaagggaatggccctgaagtgctgatcagat
gtctttccagaggatccaaccaattttaatatagccaccatcaatgcatgagactttgtagttcagggaaggca
ggcctggttttaaaaaatcatttccccctctctagcatttttctgatgtgatccttaagatttacttttagttt
tcccaggtctcattggcatgtatgctgttagggatgggtctaaaatttaatttttcttcacattcatatcatg
tcatcccagtgattatttaataaataatcacttgattaaatagtgattccttttctagttatttttgggaca
tttattaaaacctggataggttggtcctcatgctgtattcccagcactttgggaggctgaggtgggggattg
cttgagactaggaggttcaacaccagcctgggcagcatagcaagactccatctctataaaaaataaggaaatta
gtcaggcatggtzggtaacttgctggaggtcccagctacttggaaggctgaggtaggagaattgcttgagtc
agggtgcaaggctgcagtgagctatgaccatactactgtactccagcctgggcaacagagtgaactctgt
ctgaaaaaaaaaaaaaaaaaaaaaaaaaaaaazzzzzzaaaaaaaaaagatgtgtaggagcaattttggagtat
tcatttgggtcatttgatatgtagtttttagttttggtgctgatagagccagaatgtaccctgaatttgatga
acattctgatatatgggggagctcattgtccccacttaccttttgcctctcagaatatcttttgatattt
ttatctgttttttccccattgaatgttattaccttatcaagctcaaaaaagtagccctatcgctattttaagt
tcagttgtgttaaatctataaattagcttgggaaatttggtatattaatgaactcatgaagaagcagagttt
agctctccttaattctcatcttctttattttatcctactacagttctgtggttttcttttatgtaagaagca
catgztttggctaagttaatgcctaggtttttttgtttatgtgtccattctcactgtggatagtattctctt
ttccccacattatattaatttaactggttttcagagactaatagcaatgctatttttaggagaatttacct
tggttctgattaacttaccataacttgcaaatcatttgcagctttttagtttaactttgtgagttctcttaga
tttacgaccatgccagaaacagaaaggatattttcatctcttcttctgtatgtttattcttctgtttctct
ttttttatccccattatattctcaagaatctctcaataactaagaaatagcgacttcatttttcagcgzcg
agtgcattattttggctaccatgattcagaagcctcttgccctaacttgaattaacgtggctaaatctcaagaatttaagagc
accgtgactgtgtcctcaggctaggaggggaattgggttcacagagtgactggattgtggtctatgaacttc
ggcagccagcagcaaaagtcaggcatgaataatcaagtgacagtgaaacatctgtagtggtgggagatgttgg
cataactatgaatgatgattcaagagtgggttgatgcatattgaataacatgatgataagtagtagctctg

tgctaagccttctatgtgaaatacatttaattctcataataactctagagcagtggttctcgaccggggccg
gttatccccctacccacccacccctcaccctccaccaggacataacatctggagatatttttggtgtc
acaatcctgggaatgtatgtgctgatatttagaggttgaggtcagggatgctgctgaactctgtagaattca
taggagagtcctctcacaacacctatctggcccaaatgtcagtagggctactatcaagaaaatctgctctag
cagtgctgctcatattatccccatgttgaaatagcaagatgggaagtgcagagtggtgcttcggtactctt
ggagcagctttgactttggtgagaaacgccttttaaaaacaatgtttcttcccatcttccacccccatgggg
aggtgtgggggttgggtgggttaggcaccaaagcaagatttagaagagttttctgtaggaatttataatggtaa
aggatcaacttcatttccaagctatttatgagggtttatgtttaggaaaagtgcagcttagagaaggagg
agaaatctgattttattaatgagtgtagccataatggcataatcctggcagaagtcaactttggtttctagag
ggaggctattatgaaaagaaatacctggaacattccccctgggtttggaagggtgagttctaggttcaatgatg
ggaagaatttttagaggtccaagataaaaaggggcaagattaaattttgtctctcatgagttctctgggtcagg
tggtgtgaactttgcagacagtcctcttaattcactcatacatgctagtctcccagctcagcaagggtttg
agagagcaggtgtctgtatgctctggtaagtgaaggcaagtgcataaggaggttgggggtccataatggcga
agagaaggagcccttcagtcagagtggtttgaatcttggtctgcccatttgccaatcttgaccattgggc
agtgtattaactctttgaatctcagcttccctctctgtaaaaatgtgtataacaagagtagtaattggattgt
ttgatgattaaatgagttaatgtgtataaaagcactcacaacctggtagatagtaagaccttccattattat
tatcatcatcaatttttttaacctcttttctgctgcttacctcaccagcttcagctgtccaaatgg
cttgtaagattttttgtttgcccccttgctgtcagttggcatggggaagatccattcattttttcagtcac
caacataattttgagcatctgctgccccacaggtacatctagatatgggggtgcagagatatccaggaacataa
gctctgatttaattgggtcagatcagtgctcagcagggctggcaagtgctaggtttcttttaagtggcatatc
ttaaagggtatatgtcctzaacatagctttgtgatggcagcatgatgggtacaaaagcacacacttaagt
tcagtagatctgggttcaaacatttggtgcagtttcttatggctcgtaactgttcaaacctcagtttctca
cttctaaaacggtaatgatacaacctacctcacagggttattatgaattaaatactggagatgagatacaca
aaacgtcttgagztacacagtagctgcccataattgggtgtaagtattataaatctacaagctgtgaattaa
ttttacctctctggatcctgzttgatattttctagaccattccacctagtgggggccatttctacctgagtca
cccggtggtgtcaaatagaatgtcatgtggcctcctgagttgggtagaattgggtgctcatctcaacccgct
actgactatctctgtgatttaccttctcagccttagccttgctacatataaaaatcaagacaataatggt
tctatctcacagggttgctcctgaggattaaattaagtaattataaaaatgtgccttgcatatattgggc
cctaaataaacagtagctactatttatccttaaagtacaaatggtagtttcagagcttcaaggctgatggct
atttatcttactcatactctttgttttagcttcattttttccccctaatttcattagwattttcttttctct
tt
cgatcttggtcaccaccaacctctgtctcctgggttcaaacagttctcctgctcagcctcccgagtagctg
ggattacaggctcccgccacccatgcccagctatttttttgattttcagtagagatgggggttcaccctttt
gaccaggctggtcttgaactcctgacctcatgatcaaccacctcagcctcccaaagtgtgggtttacag
tgtgagccaccacgcccggcctcataagtttttctaaattttttacagtcagtcattttaaaaggaaagt
tgtattcctgtctttgttaataatttataagtgattttattcagctacaagcttggaatggcatataattttg
tattctgcttttttcaacttaataattacatggctaatgatttctgtgtttcataaacattattctgatgatgg
catgatataattgtgagtcagtcacataattgaatcatttccctattgtcatgcaattaaagtgtttcca
atatatttgcaattataatgtttcaatgaatgaataactttatgcataatagctttttgatatttaagttcag
tttctaggatgaattttccaggaatagtaattgggcataatgggataaacatgactcttgaatacgtattgtt
aacattgctttcccaaagggtcactgatatttatattccgtgttcattatcttttaaccagctcatttac
tcaccaaacatttttaagccattatcatgtggttaggcttagtaagaagaaagtgacctaaaggagaaagt
tatataataatagggtccctggtgtaccaagtgtgatacagacacaaagtacctggggaaattgagatgag
ggagtctgggtcagctgggagaaaagttcattttcatagagtcaggtttgttctttggcagaaagaaaa
ttgctttcttccccacccccacccccagctttattgaggtataattgacaaaataaaaattgtatatcttta
gatatgcaatgtgatataatgtatatctcaacttaaaaaataagctacagaataaaaagggtgtttgctatt
aaaaaaaaaagaaaagggtgaatgtcattcccaagcttggaatttgagtagttgtgctctttgggattattt
acagaaatatttagcaagaccagccccatctttggtcttgagtactccactgtcagcatgctttcttcagag
agggatccatttgctttatttttcttctgttgccgtctatgcaactattcttgatagttttatggta
acagtggtttttttgttccatgagatzaatttatacatgctcattgtggaaaatttagaaaagacaggaaagt
attaaaazacatcacttttttttttttttttttttttttttttttttaagzagacagagcttgcctgtgcgc
caggccggagtgagtgccgtgatctcagctcacagcaacctccgcttcccagggttaagtgttctcctgc
ctcagcctcccaagtagctgggagtagcaggtacaggtacacacccagcgggctaatgtttgtatttttagtagag
atgggggtttaccatgttggccaggctggctcctcaactcctgacctcaggtgatccgctgcttggcctcg
caaagttctgggattataggcaggagccactgcgccagccacacctacgttcttatcatcctagtacatcca
ctgtcattatcttctgtatttcttctgcccagctctcactctgatcatgcagtgccgtgatcatgcagtga
tctcggtcactgcaacctaggccttctgggttcagtagtattctcctgcttagcctcctgggttcaagtga

ttctcttgccctggcctcccaagtagctgggattacaggcatcacccccatgccatcctaatttttgtatt
tttagtagacacagcgtttcactaaaattttgtatttttagtagagatggggtttcaccatgttggccagggc
tggtctccaactcctgacctcaggtgatccgcctgccttgccctcacaaagtgattacaggcatgagccact
gcatccatcgccaaaaagatttttttaaagagtttaattgtagaaccatatcaaaggtctttggaaataaaaa
acagtttttttaaaaatatcagaaataaaacaacaaataaataaaaaacaccccaaaacaatctgaa
gcacgagcacctagcagaaaggttcaattatgatctattcatagagtggaaatcaagtagacattacagga
catgttttaagattatattttatgtcatgggaaatgctctccagtagatgttaaatgaaaaaacagaata
caaaagtatatatgtgcatagctcaatattgtagagaaaaatatttttatgtatgcatgaaaaagac
aaaagatgttaacagagatccattgttacttcagtttactagggattgtctctgggaggtaggattaagggtg
atttatattttaccttttttaacttttctgtatttttttattttcaaattttccataaaaaataaaggacttg
aagatcaagaaaaaattttctgctttggctcagtgacgtggctcacgcctgtaatcccagcagtttgggagcc
ctaggggagaggatcacttgaacccaagagtttgacgtccagtgagctatgatctccggatcgtaccgcct
ggacgatggagcaagaccctgtctcaaaaaaaaaaattcttgccttttttttttggttgttttgagacgga
gtctctctctgttgccccagctggagtacagtggcacaaatctcagctcacccgaacctctgcctcctgggtt
caagcgattctcttgccctcagcctcccaagtacctgggattccatgcacccaccactatgccagctacttt
tttgtattttcagtagagacaggggtttcaccatgttggccaggtcgtctcgaattcctgacctcagctgat
ccaccggccttgccctcccaagtgctgggattacaggcatgagccactgtgcccagcccaatcttttgctt
tttttaaaaaaagaagacaaaaagggattttataccagtagttatcttggtgtgtgactctgaagccacagt
tgtaagttataattactctgaaacacaaggccctgtgactctttgggctctttgggtgtttatcttgattac
aacgttggaatatagaaatgaaaggaatgggagaggtgatagacttcaggcagtgtaactagttgtctgaac
actactggctcaattatattgtgtctagtgatttccatcttgcctctgctaatttatcgccctggtaactc
actgaggcaggggttttctttggagaaacctcattgttttaaccagtgatcatgcttgtttagaagttcaa
tgatcttttttaactcatcggaagatgatgaccagacctggacagatggggaaggactttgcactctctct
ttacagtcctgagtgcacacaggtcaatatggaactatgtgtgaattttcattgtctttgagagccctcttc
tctgccccatagggagcagctttgtgtgcaattagaggagcaaggggtgtgtgtatttagcacagcaggttg
gcctggctcctctcctctcaacatagtcaccacatacctggcactatgctaaggctgggaatgcagacagatg
gggtgcctgctttcagagtgtcattgtgtgaggaagccagcaacagaaacagatgatttcaggagctccag
gaaatgctacaggaggagtgtgcctgggttactggagtgcacaggaggagggttctagctcaggctgag
attttagtaaaggaaattatgccacgatgaatcctgaagaatgaatagaagtgaaccagataaaagcacgata
ggaagcatcttcccttacctaagggaagacacagaggtatatggaatggtatgttaaagggttgggactcca
aacagttctgttaaagcttagagagtggtgggagagactggagaagttgattaattagtaaatgaagttgtc
tgtggatttcccagatcccagtggttgatgatccatatttttttaaaattacagtggttctatcttattt
cccactcag**TGTCAGCTGCTGCTGGAAGTGGCCTGGCCTCTATTTATCTTCTGATCCTGATCTCTGTTCCG**
CTGAGCTACCCACCTATGAACAACATGAATgtaagtaactgtggatgttgccctgagactcaccaatggcag
ggaaaaatccaggcaattaacgtgggctaaattggacttttccaaagatgctgtctttgggaaacatcacaca
tgctttggatcagzaaaacctaggcttctaatttgttgataaggcatgaactcaggagactgttttcagtc
tagtgaatgggtgataattgtaattataacagtagacaacatctctttacacattttaaatcatgaaaatag
aataaccttactgataaattttagaaagtggtgattaaaagcacatttaagataatgccttaacacctagtct
tttccatatgcatgatgtcttaatcacacattgcaaatcatggaacacagaattttaagcagcatttgtgta
gaacttctcagttttactaatattattttattttatttctcataacaaccttgaatagaactcagatcatctg
tcaatcatgtattttgataacagcctttacagtgagcatagaaaatacagtagtggttaacaacacaggctc
cagatgtcaggttatctgggtatgaattctgggtgtcagcattcactaagcatatgaccttggacaagtgatt
taagtttcttttaaacagagaatagtaatacctacctcatattattattgtcagtgatcatcttacaatca
cagcttttctcttagggctgggctcagtggggtgattgacactgcagaaatggccagatctaaggatcaac
atttacgtagctgggaaatgtagctgggacttcagtttactgccttagtgatttttcctaccactaagcag
ctcagtcacataccctacagagacccacaagcttatgagatactgttcttccaggaaagcagtgggggccaggg
ccaccttttaattgtgtttcttgccctgggtcccactctttctcacaatatatagcaacagttatttacttgc
gattttctaatgcacatcacacatagtcataataaacacacacacacacacacacacacacacacccctc
aagaaacattttctgagacgtgatttctgatttcatcaaaaaagaaaagagcggggccaggcacagtgggaa
gtcaaggtgggtggatcacttgaggtcaggagtttgaaaccagcctggccaacacgggtggaacctcgtctct
actaaaaatacaaaaattagccaggcgtgggtggcgacacctgtaatcccagctactggggaggctgagga
ggagaattgcttcaacctgcgaggctgaggttgagtgagccagattgcgccattgcacctcagcctggggc
aacagagtgagactctgtctcaaaaaaaaaaaaaaaaaaaaaaaggcataaaactgaaatttatatgcaatttat
atgcctgtgagataattttctgttttctcttttggaaaccccaagagatttttttgattgatgagcaaatat
tttagattttatttaagcattatgccaaagcaccactgaagtataagtttcaagggcaaacctcagttttttca
tctactagacgaatgattttctggaatgattacaagcaggcaazgatgggtgzztagtggaatagcaaatgt
cttcggcatcagacaagttgggggtttgtttgtatcctgcctctgccttccaccgaggttgtgatcttgggca

gattgttgagttttaacctagattcctctgactccagatcataaattttcagaaaagttctgaaattcttgt
atatactgagtggttaaatgagacttttcttacatctatgcacttctttgtttgtttgtttgagatggtctt
gctctgttgcccagactggagtgagtagtgcacatctccgctcactacaatgtctgcctcccaggttccagt
gagcctcctgcctcagcctccaaatagctgagactacaggcagtgccaccacgtccggctaatttttgta
tttttagtagagacagggttttgccatgttgaccacactgggtctgaactcctggcctcaggtgattcgccc
gcctcagcctcccaaagtgtgggattacaggcagtgagccaccatgcccgccatatccatgcacttcttgc
aaccttaccttcttttctcatcaccctccagggacctaagtggagagcagagttaaaagttaaggtgaaac
ttggagaggtgtcttgtccctaggaacaaaggactgggttgaaattctctgtaaattctccccaggttcaaac
cagagttatcaaggtcttaaaaacttccctgggtcctgagagccattatattatttacttgtcttccctgta
caccactgcctagtcctgactcctacttttgtttgcaaataaggatggggcacaacgtacaaggaagggcctt
tgccacccctgctaagggataacctgaaataccttcaccatcactgccctgtgctgcttttccctatgcc
gtctgtctacagtgccagtgctcctggcattgaaaggggagaatcttttgggtcctttgagatatttgggttg
gttacataaatctccctgaatgaagagcagctgacttaggcaaggggccttgtttgggttttccctgaactat
taacaggaagatagggagattaactgtgtaaattgtcaataggccagagtccttgcagaggggtggccacagt
gatcagatcttatcacatccttgctttgggtgttgccctctctggttggagtaggtagaaaagaaagaaag
accctatattgaaatgcaaagtgcagcaagtcctgactttggattaaacttctcagcccatttgcagtaaaaat
aaaaagatgaataaaaacaaggttcccactttggagggaggtggttagctgtgagatggaagaggtgttccctgc
tgggcaacagcagagtaagtgtgggtagattcactcccacagtgcctggaaaatcctcataggtcattt
gttgagtctttgtcctacaccaggcactctgcaaaaacgctttgctgcaaggtctcatgcatgctcacca
cagctctgtgaagttaattgtacttttatcaccattttacagatgagaaaactgagggtaggggtcaatga
cttggctaaagtcactgcttagcaagctgcaggactggatgtgaattccaattgggttgactccaaagcct
gtgaagctacttgttcttccaccactagagctgtgggtcttgataactgtgaactcttttggggtcacaaat
agccctgagaatatgatagaagcaggagctctggccttctgtccatacctgaacaggtccttgggttaaga
gccccctgctccagggcctattaatcttgatcctcataagcagcatccatgtattacggccgcaaaccaact
gtgccagaccgaatcctaggaccaagccccaaatatgtcccatcatccttttggtaagaagctcattgtaaga
aagaaagaggagagcaagaggatgacctagtgcaggggcctcattgttttaattagtacaaaaacaacaat
aataacaacaaaacccccgaagcttcacagatgacatcagaccccaagcctgtgtgttttccaggtgccctt
gaggagctttgtagctggcagaggaggtgaaactgacaaatgttggcagatggaggagagtaccagagggg
tttgagatgagctaaattccaatctaaccgcaggtgtgaggaagaggcttggattgggaccatggagatggg
gggtctactcccagtcacgccagctgactttgagagtggttcttctgagtcactttatcttattttat
ttttatttttttgaaatggagtttgcgtcttctgcgccaggctggagtgaatggcgcatcttggctcact
gcaacctccccctcctgagttcaagcatttctcctgcctcagcctccagagtacctgggattacaggcgct
gccaccaagcccatcgaatttttgtatgcttagtagagacaggggttctgcctatgttggccaggggtggtcttg
aactcctgacctcaggtgatccgccaccttggcctcccaaagtgtgggattacaggcgagccactgtg
cccagcccacttcatcttaccgtagttaacctccttagagtagtaaaaaataggcttagggcatccccagtc
ccctctatgtctgagagctgaggtggctgtcaaagaggaaactaaggatgccagggacttctgcttaggac
gcatgtgaggccaagtttccacctgagagtcagttttattcagaagagacaggtctctgggatgtggggaat
gggacggacagacttggcatggaagcattgtataaatggagcctcaaaatcgcttcagggaattaatgtttct
cctgtgtttttctactcctcgatttcaacag**GCCATTTTCCAAATAAAGCCATGCCCTCTGCAGGAACACT**
TCCTTGGGTTTCAGGGGATTATCTGTAATGCCAACAAACCCTGTTTCCGTTACCCGACTCCTGGGGAGGCTCC
CGGAGTTGTTGGAAACTTTAACAAATCCATgtaagtatcagatcaggttttcttccaaacttgtcagttaa
tcttttcttcttcttcttctgtcctctggagaattttgaatggctggatttaagtgaagttgtttttgtaaat
gcttgtgtgatagagctgcagaatgaggggaagggagaattttggagaatttggggatatttggggatccat
cacctcgagtatttatcatttctgtatgttgtgaacatttcaagtcctgtctgctagctattttggaatata
ctatatgttgttaatgatcatgcagcagacgtgcacatgaatgggctggctctaggagctagagggtagg
ggctggcacaagatgcatgctggaaggtccttgcccataagaagctgacagccaaggctaggggagttct
gtcttctctgcatcaggtcacctctctcacctctgtcactgccccatcagactacaatgtctgcaggtcttt
ctccccctgagtgtagctcctgagcaaaagcaggatgctgccccctcccccttcttcttctggctcctggctt
cagtgccctggacataagtatgggcataataagtgtcccccaaatgagacattgaggattcttcaaatgcaca
ggaccgtgatgtgagttaggacggagtaaggacgatgggatgtggctcaggacaactcctgaggaagctgcag
ctgcccacgcagggccacactgtcatgttcatggaccctagactggcttcttgtagcctccatggggcccttc
catacacaataattaaaaattatatttcatgactgzcattgggtataaagatgaatataatccagaccagatt
catgattattcatacatttttagtgtattaacttttaattctgcttttaaaataaattaaaacatttcaata
tgcccttaagagtatcccaggcccaggccactgagcctactgtgggtcatggataagztggccctgggggc
atgtgtgtgcatgcatgtgtgtgcacatgcatgatgagccgggccttgaaggggtggaagatttgggtgtgt
agaccaatggagaaaggcatttggggcagtgatgatgggtgggggaggggaacatgggtgatgaatggagctgg

gtgtggggagccatgggagtggttagggccagcctgtggaggacctgggagccaggctgagttctatgcac
ttggcagtcacttctgtaaagcagcagaggcagttggcctagctaaagcctttcgcttttcttgcaccctt
tacag**TGTGGCTCGCCTGTTCTCAGATGCTCGGAGGCTTCTTTTATACAGCCAGAAAGACACCAGCATGAAG**
GACATGCGCAAAGTTCTGAGAACATTACAGCAGATCAAGAAATCCAGCTCAAgtaagtaaaaaccttctctg
catccgtttataattggaaattgacctgcaccagggaaagagagtagcccagggtgtctggggcttgttccca
ttagatcttccccaaggggtttttctccttggtggctggcctgtggggccctctccaggaggcatgtgtga
agaaactaggggagctggttggcacagacagtgatgtactaatcttctctgggaagacagaagaaaagtc
cagggaagaatactacagacttggccttagggacagctaggggtgcagattgctgccaactgcatttttct
gaagttggccatattggttgcagtgaaatggatttatagacagagtatttctgtgcatataagagcaattacag
ttgtaagttgatattgataagtgaaagttaagcacttcttctaaaaagagaatgcaattcattttcccta
atcatttcaatttagtctgatzggcatttgaacttgttgtctttaaagtgaaatctttacctctgatctg
gtaagtatccaggcaatttcttgtgtgccaccaggaggtatctggggagtgggcattttctgactgaggca
ttggctgccatagcatcagagcagccttccaggcagtggcctggcaaggggacagaggtggtgggagcagc
tggctgagtgagccagtaattggcatgtgcatggtctgtagagaatgtagaagcaataatgaagccgataaa
agctggtctgcattttattattatcatgcgcgggtggttctaaacaatgtcagtgataaattactcctccc
atcatggaccaatggctgccactgctccagggaagtgttttattcctgttgggtgtttagggagggatgga
gttggctggccttctgtgaaaggcctaccagtttgttttctatttggcaaaagaagaatgataaagttyta
gagtttaaaccagactcagatttgagtttttttttttttttttttttttttttttttttaaggctacagaactgtg
ctttccttgggcagtaaaagaggcaatgggacatgtgggacctgatzzgacaaazagggaaazzzgzzzzczaz
zzazzzzzzctgtcttaggggtggcatgggagggaggtgctgcttcacagcagagagaggtatggctgtgctt
ggagtgctccacttagacaactcctggctgtgcagccaggccatcgagatgctgttcttgcactgcaggtc
ctggtcttgcacatggatgttcttctggtgcaggagacagaaaggtagcaacaacccctgatcaaaagcctc
agtccttcttatttactggagagccccctgctgattgaccagaggcacagctggggatatttctttacctc
tgtagcaagagacagcgtggtgcagaggaaagtgtcagcatacattaccctgtggctgcatgactttgtgaa
taggttaggttagcaccctttcagccacttcttcttacctgttaatgagataaaacatgtaattgcttaaaaa
cagtatttggcacataggaagcacttagtgaatatgaattatgatttttttggagtggtgacatctcaacc
aagccatttaacccctcagccttzaacttctcactataaaatagcagctaacttgaatgtaaactataa
aacctaattgtagtatctggcacatagtagattcccaataaatgagagccagttattcttctaaagacagtgat
gcatttctgagcacctggccttgttcttctgccttgcaatttatgcagcagttgaaatagactggctgatgg
gggtaagttgtcaagcagacttctgatcttagtgaggagactgccttaaaacaacactaatttctttt
cttttcttttcttttcttttaagacagagcctcgtctgtcaccaggctggagtgagtgccgagctcttg
ctcactgcagcctctgcctcctgggttcaagcgatttctcctgcctcagcctcctgagtagctgggactatag
gcatgtctccacatgcccaactaatttttgtatttttagtagagatgagatttcaccatgttggccaggatg
gtctcgatctcctaacctcgtgatctgcccgcctaggccttccaaagtgtgggattacaggcatgagccac
catgcctggccttctttgagaagctggagacatgagtttaagtggtgaagaagccaaatctgtatctaaaaac
cctacagtagtggtgcagagctctgaggagagaaggtcccttagattttgagtgattattatgtcagtgctt
gttttacatctctctgttccagcagtatgtccccttttctgccttgacgtgttctttaaattctttcttct
tttgcttgtcttgcagcacaaaaacaggcttcagtatagggggaaatgcacagaaacactgccttttctaca
ggaaatcagtaactttttactgattttgtttttatttacttattttatttgttttaaatttatttttagtttt
tttttttttagagacagggctctcttctgttaccaggctagagtgagtggtgcccttagagctcactga
gctcactgcagcctcgaattcctgggctcaagtgtctcctgccttagcctcccgaagggtgggatgaca
ggcatgagccactgcacctggccaacttttctgtgattgcgaatagcactcttgtcaatttcggagagaagc
tgagactggcatatgtcagtatggatccccacttagagacctgtgtttatctgcactgacaccccatcacag
catgatgagcttggccctcctgtgctgtctctcccagggtgggaggatccttgaagctgatctggttggga
gctttgtcctcattcacctcctttaccacacacccaccttcccaggggggggtatctaccactcactaagtag
cccattctgggtgttgacagctctaattgttagaaaatatccaccacctgttatgctttctagagaacaag
tctaattctgttttcttgaatagtcgaagacagctctcatgtttttcttccctgttttcccaaagtcca
tgatttttttaggcaaatggcctccttctcttatggaatgttttcttcccttctgctctcctctg
gttgtgtttcagtatgtctgtgtgttcttgaagtttactggaaattatgaaagtattctggcacagaggag
gaagggttttgcctccctgttctgagtgctacatttccgttaaatgcagctctgagattgtattaggcatt
ttggcattcacgtcaccttgttgactcatattccatgtgcactcaaacaaaattgtgattattttaaatagg
cagaattgcaagttacgtgttctccatttcttgtgtattgttggctttttgaaactaaagggaataatgtc
tttttctgtttttacatgttttagattccctatgctatcctctccaaaaccatttttagattctgatttt
gccatgtattatatctgatactccttctcgtcatctagagatgtgataaacaactctcttggcctcattc
cagtcacatgataactgtgggacaaagacttgaagcttggatcagtcagtgaggactaaccacccctgtaga
cccttttttctcactataaaatagcagctaacttgaatgtaaactataaaacctaattgtagtatctggc
acatagtagattcccaataaatgagagccagttattcttctaaagacagtgatgcatttctgagcacctggcct

tgttcttctgccttgcaatztatgcagcagttgaaatagactggctgatgggggtaagttgtcaagcagact
ttctgatcttagtggaggagactgccttaaaacaccactaatttcctttttcttttcttttctttt
ttttttaagacagagcctcgctctgtczaccaggctggagtgcagtggcgcaatcttggtcactgcagcct
ctgcctcctgggttcaagtgattcttgattctgtagacactaccactcaggcctatattgtaatcagtgcgtg
ggccactgggctcctgcttctgtgatccagttgggaagtttatcttggtcttcccttcagcttgcattctgct
aaattcgctggactatacacaggtgattttagatatggggatctctactcaaatactctcatgatttcctt
ggctagagcatcattttatttccacttattggaagagaccttagagaccagtttagttcatttatagataaat
tagttgattctgtcattcaacccttatatatgtagcgtctcctatatgcgaatcactgttctaagtgccgag
acacagaggtgtccaaaacaaatatggccctctccatatggaatttctattctagagaagaatctgacca
gaagggggaagtgactgtcccaagtctacacaaccacagaagggatattctgggaataaatcacggctaacc
ccccctgctgctccaggcagttctccttctacagtgccttattgtgctgttttaataatctttcaactggga
agaactcccatttcaggaattaagccgtggacaaatcttttaattatccttgaaatcatcctaataagaaat
ccaaggaggaagtcttacaggggtgcctcacccacttttctcatcactggaacttttttagacattttattatt
ttcttccataaaccagagtacaggcacacaagttgagtggtgtggtggctaaattaattaatgtttgcaaggc
agtgtgagaagcattcattcatcttaaatacctatggtgactgcaactcagatgtaaaaattggataaatcc
tcagaaaccctagggaaagtgcacatgtctgtattttgtctctgtgagatacagactggcagagataagtggt
tctctgggtgagttttgtgtggtatctgggatgattttaggcagtagctggatgagaacttttaattttaac
ccacatccaattgcaatttcatggaaattattgcttaggaggatgttcaacaggaaaaatataattaagtt
aattcaaaagaaacattttctgtgaatatggttaaaacttgtgagagtgtttgtaaatgattgaagattgga
aaacatttggataaagagtgagtggtgggttttattagattcattttgggaagaaatccatgctgcaccc
ctcatgaagtgtagaactttgggcatgtgttgattcttctggcccagagtttacctgaagattagctgcct
gagggctactgagcattaaattagatgatgtctgtggatgactgatagtgaaagctcatagcccacagttgac
acataataaattcgagttgctttgcttcccttctgttctggtgactgtttggcctttgccacttgttcg
gcctctctgggcttaagtttctttgccttgaaattggaactttctttggtgaaacaaccagaaaatgcttc
agcccagaaaacttggtcagtaacttggatgggggatcacctgggactacccaaggatgtgggctgtctgctag
actatagcccttgagggcaaggtgggctgctcatggttccatcctaagcccagcacagtagtggtg
gcatggtgagccattagtgaaatcttctgtggaatgaaggtgggagaaaaataaaataacctgtacttcacaggg
attgtgaggggtcaagtaaaagtgtttaaaaaattgtattatagtttattcccttgtgttagcccagggtc
aacagagcctacgaataataatgatgcagaagttcttcaaaaagtcttggccttcttctttcacaaaatt
gccccccagagctttctggaagggcagccatgaaccagaggcctaagtagatttactgggaagctaaaaa
tatttactttatttttcatagctcctttcaaggctcctctctgggggtcttagcaatatgtttacacagtgg
atgtttttgtaaggtttgcaaaagtaagatttttaaaaaatactatcttgttttaaaaagagagcccctac
caacttgtgtcagcctcaggccccccacctgcacatctgctcctgcccagggcaggtggggcaagaagcactgt
ccccttccaaagcttccctccttgccttgagtcacctcactcccactccaagccacctgcccacgtctgt
gccccctctctggtgaatctggcattcttaggtgggctgagaagcagactggcccaagctaaggcctttctg
atgggggtgtgtgctgctgagaatcatgactgggtgggagaggaggtgacctttzgtgtcttatttttac
tgtgtatttcttctcagctacttaaaatgtattgttagtgatacctaaggttcatttagcctgcttccct
actgaacatttccgctcaggcatccacttgggtcccaaggcctgctcctctcccatattctgaaatctggact
acagactctcattcaactccaggttgcaactgtggacacagtcctccttgagcaggtacctccttgcaagtgg
ttgggacgtcctacttggctcatagttgggaagtgcacatgctggagctgaagcctcttgccctccgggat
aggcgctcctcacatccccctctgagaagttccccagcttccctctgttccccgtttccacacttagcgaggc
tcttgctccactgctacatcccccatagccagtcctcagccttgccattgcttatgctggctctggaacaat
tcctagacttgtggggcatctggggaagtctccatcttttttagctggcatgaccaagtgggtgtgggca
gggctgtggattctatggtgtggctggaagccaggtagcctctctctactgtacatggaactcagcaacttc
tgagtcaagcaagatcttagctctgcaggtgtcttgccctgtccaaagttatggccacaccagtacctttta
actctagaagcccagtaagtgttttgtgggaccgcaaagatcattttctagacctgctgaatatgcctagaa
cgggtagggatggctttcacgctgttccctagggtgacaagtccacgctttctgggggtacatacacaccgc
ggtccctgtgaatggcactctccatgagaactgtgatgattgagttgaatagtgcacagcctacatggttc
tctgccatggcctggagttccttatcttgcttctccagtgaggactagggtgcaactggcctactttggc
tccctgacttgggggattctgaaataccttttttttaaggttgtggagctctctgaagcttataaggattt
tgccaggaaaagataaagaaatatctttgggcattttgtcactgtgctggagatgaacctttggaggacata
tcaccttggtgaggtcaaggggcgaaaggacaggaactggcagagagatccggggcagcagcctgccatcc
cgactgagtatggagtttctctcctcctcagctggacttttgggtggagtcagtggtcagctgccacttc
ccttatggttcatggcatgaatctggcttgttaggcctttcttttttttttttttttttttttttttttttt
ctctgttggccaggatggagtgcagtgagtgatgtggtgtgatctcggtcactgcgaccaccgcctcctgg
attcaagcgattctcctgcctcagcctcctgagaagctgggattacaggcgcatgccacaacaccctgctaa
ttttttaattttttatt

tcctatgtctacttctttctacacagacacagcaacaatctgatttctctatcttttccccacatttcccc
ttttctattcgacaaaactgccattgtcatcatggcccatctcaatgagctgctgggtacacctcccagat
ggggcgggcgccgggtagaggggctcctcacttcccagaagggcgccgggcagagggcgccccccacctcc
cggacggggcggtggccgggggggggctgccccccacctccctcccgacgggggtggctggccggggcgggg
gctggccccacctccctcccgacggggcggtggctgcccctgctaatzttttgtattttcagtagagatg
gggttttaccatgttggtcaggctggctctcgaactczzczztgacctgttgatccacctgzzcctcazgc
zzctcccaaagtgttgagatztacaggcatgagccactgcgcccggctctgttztttttgtttgtttgtt
ttttgtttttgttttgagacggagtcttactctgttgcccaggctggagtgagtgaggcgccacgatcttg
gttcaactgcaacctctgcctcccgzgggttczaaacaattcttctcctgcztzcagcctcccgaaatagctggga
ttacazggcacttaccacczagggcctggctaattttttagtagzzazgacgggggttttgczcat
gttggccatttgaactcctgaccatctttzgaactcctgaccatgtttgczcatgttgaactcctgacttca
gzgtgcgttggccztcccaaagtgtgggattacaggtgtgagccaccatgccagcctgtttggcctttct
gatatzgzczctcztgactaatctttttggaaattzagzztccccagggttatactggattttacttaggg
aaaagggtcatgcztctctggctgtcagzatttactgatagtaactaaggzactcagztgggggtgzzgzac
ctttgattczztzzggtttgatttttgaaaatzcaaaaagacgtgagctccaggggagcaggggtggctttgggtg
azcatggcaagatagttggctgtggczagggagttaggggaagtgggtagaaaattaacatcttgtazaata
ttzzczcctgggaaatatacctctgtgttaagagaacagacttggcagzczzazzzzzzzzzggcc
taggttcaaactcttggcttgatgcttatcagctgtgttaaccttggataattccatacatctctgtgcctcag
tttctcaaattggaataacaatagtagtacctcctcaggactattgtggcaaattaatggacgaataaggggaa
gcacttagtagtgcctggccagcataggtaccaggcttgttcttaagctcactgcatttttacaatcat
cataaattgcaggggatacacacatgaaggagccgaagttcagagaggccaagtaacttgcctaagaaagca
cagctggcaaggggcagtaataggaccagaattccagtttctcgtgctctgttgttttatatcctaagag
agcagctctgagtagccagaagcttccctaaagtacaggacatggggcatgggctggctgggatgagaaag
gagacaagagggcttctgaaagaaatgccagattcactccacttctggcttcaggcaccgatggaatgttt
cccaaggcccatctagaaagaacatcctgtgactcacagccacttctcattttctgtctgaacccctcacc
cattcaggcagctgctaaagttagggttacagcctcagactatattttctgtccttgtggaacccctcagtg
tcatcttgttgggagatctgggtgatacatgtgtcaacattatgtcatcaaaatggaaattctttgaaatctt
taggtgattgcaattcacgttctgtatgtatgcacttgtcaaaagttttgatttgaggccttagaattttat
atttggaaccccttccactaccatgagttttccagacctgtcaaagccaggctgcatctcagaaaccagggt
ctttgatttccatccagggcaagggcctggggccagctgggctgtaagcagggtgggggtggggagcaacgct
gcaactgcaatgttgaaatattacttgaactaaatcaaatcaagatcagcttactcagacaagaatagaaa
acacaattgcattcgattacagaatagtgtgtatccccacaaatatcagactgccttataaaagttttgaat
tgttaacatcaagaacagtggttgctgtctcctgcttttccagcataagggttatttattctgtgggtggca
aagagcaatttgggagtcagtttgtttctcattgaaagctttccatttctggctccttctgctactgttg
attgaggcaccaaaaggcaatctcagtgcgacactattcaacagactaagttgcaccggataatgatccat
tttacatttttcatatattattacattgaaggcttcaaacagcactagcaggggcaaatgggtattattatc
tccatttcatgtgaggaacccgaggtcgaagggtaattgtctgttgcgaagattaaagagtaaaagtgtg
tacgttgaatcggggtctgactcctaggttagcattttctccccacactatgctgccatgttgcttattcc
aacattgaagcataggtgccatccccagcttttgaggccaatatcacgatgaagcatttttaaaacatct
cattaaattgctgatatagtggaaagaaccaaagctttgcagtcгааagctgtttgggttcaaatttaccact
tgtttgttctatgacctgagcaagatattctccagatctgtttcctcatttgaaaaatgggaataataata
cgtttctttataggctgttcttaaagattctggaaataatgctaatagtgtgcctaagcttggttaaatat
gagtcacttttctgtgccacaaagcactactatgtcccttaataaattttgttaatttttaaagttagaaa
aaaattaaactatttatacattgtgtatgttaattcttccctagaccagccttaggaagaatctcatcccca
acttgtaaactcatctttttccgttctttttgtgcctggacttctcagggccctgcaggtgattctagtcc
catgttgtgtggtgtttgaagtgtctggtccctttttcagtgagagaccagctcatccttgggaactgaat
gcctcaaactctcttttcttttctcttcccttctgtttgatgtagtctttcctgtttctggactctgt
ttcttcatacttctatctcttaccttcttttcaactccttttgtcttccagctgtcctctctcatttttctg
cctctctggtcttccaggtagagttttcatctcagctatcttcttctgttttctgatgttgggtcttctgttct
ttcttctcattctgttccaggtccaaaattcatttgggtcaatgttatgtcttagtgttatcttccattcct
tctgagcttcaaagccaggctgactgtgcccttccacgccttggccagtgtagcaggacacacttccctc
tggggctgcactggctctttgggggaattgttaccattcagggcttccaaacctcattctagggacttccag
taacttctcccaccttccctcttctcagtaagacatgggtattgtcttattctgtttctgctgotagaagaa
aattctcgagactgagtaatttatacacaatagaaatttacttctcacagttctggaggctgggaaatccaa
aatcaagggtgtggcaggtttgggtgtctggtgagggctgctctctgcttctaagatgggtgccttgttgcctg
atccttaggaggggacaaacgccatttccctcaactggcagaagggactgaagcctctccctcaagcccttt
ataggggtcctcattgtcaggcctctaagcccaagccaagccatcgcatccctgtgacttgcacatatacg

gccaggcactgctggcctttggaattaaaaaccctggccctcaaggaatttctattcttctgggagcaggata
 ttacatgggtgactgczagcttgatctgtgtggtactagtggaaaacataggatgttgggcacaaaagagt
 cacaaaacgcccggcgcggtggctcatgcctctaactcctagcactttgggaggccaaggcgagtggatcacc
 tgaggtcaggagttcaagaccagcctggccaacatgatgaaaccccaactctactaaaaatacaaaaaggaa
 taatagctggcggtgatggcggtgctgtaatcccaacaactcgggaggctgaggcaggagaattgcttga
 acctggggctttggaggttgcaatgagccaagagcgggccacttcaactccagcctgggtgaaaggggtaaac
 tctgtattaaaaaaaaaaaaagtcacaaaaaagggtggccacctaacccagcttgaggatgggaaggccag
 ggcaggcttcttgaggaggtgactcatgggccatgtcatgaggatggggtaagaggagaaggatgtttca
 gacatctggcctttcttaatatcttggcctttccaccattttcctgccagcaacataggggagaagactgag
 accagcagatacaaagccactgtactctgtcctcaccctcctcttttctcccccttctactaccagatc
 tgagggttttgagaaatctcttctctaattatctgctgttcatcattgttgagaatctactatgtgcata
 gcattatgaccacaaaatgtcctggctggtgtcatcatgagaggtcttacccttctctccagccacttct
 agggatttttgactctgtccttttccagaacttggctccagtctggttgcctgcctgaagcacttacagat
 aaacctcatcttggggcagtgcttccatttactgtctccttttggcttgccttctctctgcttctt
 gaattgatttgcctctttatctccccctttcagccttgaaagttcctcgaaggcagggactgtgtccccatc
 tttctataaatggcattgttgtacattggtaaaagttgagtgaaaaaatcttttgataatcacaaattatta
 ccatatattgagcacctaccatgatgagatctctgatcttcatactttgaaatgtggatactattacct
 cacttaattggtaagaaaaatttgtgctcaggaggttaaaagtcactttctcacagtcacacagctatttcaca
 ggagagtgcagtatcaaatttaggtttttctggctcctcctcctcctcctcctcctcctcctcctcctcct
 gacttggctgcaatcacctgaaaaattttattaaattcttttttttttttttttttttttttttttttttttt
 ctgtcaccagcctggagtgcaatgggtgcatctctgctcagtgcaaccactgcctccacagttcaagtgt
 cctcctgctcagcctcccaagtagctgggattacaggcaccaccaccacactcagctaattttgatgt
 ttagtagagactgggtttcaccaggctgatctcaaactcctgacctcaagtgtctgcccacctcaccagc
 cagggttttataaattctgatgcctgggtcttaccctagcagttctcattagtgtcctcaggtgtggcctg
 ggctttgggatttttgtgaagccccagctgagctaacgggcagtcagtttgagaaccagtatactacac
 catgttgccttctgtttttcacacagggttggtgctagtgtgtgagcttaagcctacatgggtggatatccac
 gtgatcaggctgcaagttctttgtagtggaggatctttggccctcctgctgcttcccaaccagcgaacct
 accacaccatgcacgtgcacagccaagggttgggtgactgttaaatcagtcctctgggttcttatgtctcatcc
 gggggattcactgataggaggtggagcttgatcacatccagatgttctcatcttcagagcctgttaacata
 acagaagtcttataaacatgctgagcacactcactggctcggagagtgttgaggatgacctaggccccattg
 cagggttggcctgggtcgcctcccacacttctgccaccctgcttgggaaggaggagctgtttgaaagttc
 cctggtgcaatatgatgtattgtctggtctgctcaggaggaagctatggccacctagtcagtgagggtt
 agctaatacgtgtattctgtttctgttgagggtttcacagaggccttttttgacccttcattttatagataag
 gcagtggaggcacaactacatgaaatgactcgacaaataaatggatttagaaccagggtttctgactcccag
 ggtggtgctttttccatggttgtagctgattaatgtctacctttttcacaccagtcctcaactgaagacacc
 agcttacaccttcttctgtttctcccaagaacagaaagtgaccccgatgtcgtttctgttctctgggaa
 ggcagttccagtggttagaagtcctgttcaactcctgggttggtggtggtggtggtggtggtggtggtggtggt
 tcttctgtgacctggccagctaaaaggaaatctcctatgatggtactcagatacttttggaaacctgtcagc
 ctaatcccatctcctagtggttagtattcagctacccttactgggcagtaattctgtgccaggcctgatc
 tgggcattggtggtactaagaacgcataatccctatccctataggcatagtctggttagagacaacatgcaagt
 aaaacaatgttcttaaaactgtgggtccacacctccctcccccaacattaaaagtgaagggtgcttattc
 aatgtagattttaggctctgcactctagaccactatttcagaatctctggggactgggccaagaaact
 gcattttcgcagctccctaaatgaagcttaggtgctctgaggtttgacaactgcagtagagagcctaagtc
 taacagtgtagagtcacatgtgatgggaagcatcaggtaggtagcagtttgaggagcactgattctgaggg
 aactaactgggcctaagaacagctactggctgtcatgaggaataactaggagctagccatagaggggtagc
 agtgaatcatttctctagcagtgtaaatcttgcctcaatttattctgtctatataactcaatattactgaagt
 ttgcctaaagcagaatacacctggatcatacagcatttatgagagactggctgggtgctcaggccctcctgt
 tactttatctctgcagtggtgacctcttasstccgcggttaactcctgtcctcattaaagcctcacactgtag
 cccattttcagatcaaacctgttctcttctggtaaatgatttcagtttgcaaagtttgccctctagaggt
 tgcttagtgcttggccatgtgggtcagttcatgtggtcctgatgagctgggttttattctttattacaagaa
 gttaggctgttaggagagtgggttggaaggagaagaggtagacagccaaatgagatgagtcagggaaaaacta
 tactgtttcggagtcagggctctaccaagcatctggtcagaaacctctcattttggagatcaagaaatt
 gaggttcagaaagatgacatgaggtacgcaggagccaccagacacagcctccaactctagaaactaaaat
 tctggaattcttagtgctgctttttctgttttgcactggattgaagccttttctaactgtactcagaggg
 cctattatttagggagattccgtatgaaatcctttagcaatcaaatcatttaatagggtgatgggttaaat
 atatgttaatgtgttttccaaagcctggcagacctgggttttgagctttgtatcacatgttttatgtttg
 gaatgaaaatgagaccatgtctgtgaaggcactttgatatgcgtaatgcactctgccagtggtttgtcaaac

atgggtccccaggtcagcagcatcagcatcacctgtgaagtgtattctccagtcgcatccccggcccaggmsmggg
 ccmrrctczyaggytgtgsmggcsmggzssggzztgzccggccccggccccggccccggccccggca
 gtctgcatttttaacaagctctccaggtgattctgatgcatacttaagtttgagaaccattgcttgttttgca
 ttaaacaggagattagctctctgcagcttgtggaataaaagctttaaatctctccaatttttagctctgtgaaa
 aggagtggggagacaggaatgaacggactagtgccacaaagctcaggtggggtgggtgagatcatttagaa
 gagaaagaccgggcatggtggctcacgcctgtactgtcagcactttgggaggccaaggcaggttggtatcaca
 aggtcaggagtttgagaccagcctgcctatcatggtgaaacccctgtctgtactaaagataaaaaaaaaaaaa
 tttgccagtcaggtgatgcatacctgtaatcccagctactcgggaggctgaggcaggagaatctcttgaac
 ccgggaggcggggggttgcagtgagctgagattccaccattgcactccaacctaggtgacaggggtgagactcc
 gtctcaaaaataaaaaaaaaaaaaagaaaaggaaaggctgtgtgtgtgtgtatgtgtgtgtgtgtgtgtgtgtg
 tgtgtgtaacagcaccatcacactgtttgagttgaggagcacatgctgagtggtggtcaacatgttaccaga
 aagcaatattttcatgcctctcctgatatggcgatgctcccctatctcattcctgtgtgtgtgttttagccaggc
 aactgttgatcatcaatattatgataacgtttctccactgtcccattgtgcccacttttttttttttttga
 gttacttactaaataaaaaataaaacactattttctcaatag**ACTTGAAGCTTCAAGATTTCTGGTGGACAAT**
GAAACCTTCTCTGGGTTCTGTATCACAACTCTCTCTCCAAAGTCTACTGTGGACAAGATGCTGAGGGCT
GATGTCATTCTCCACAAGgtaagctgatgcctccagcttccctcagtagggctgatggcaattacgttgtgca
 gctactggaaagaaatgaataaaaccttgtccttgaatggtggtgaaggggaggaggaggtagtttgaataca
 acttcaacttaattttacttccctattcaggcaggaattgccaaccatccaggagtggaatatgcaacctgg
 cgtcatgggcccagctggttaaaaataaaattgatttctggcttatcacttggcatttgtgatgatttccctcct
 acaagggatacattttaagttgagttaaacttaaaaaatttcacagttctgaggcaataaacctggttgaag
 ggttattgatctggaggagctctgtctaaaaaaattgaggacaggagactttagacaaggggtgtatttggaga
 cttttaagaattttataaaaataagggctggacgcagtgggcactgagttgagaactgttgcttgccttgcatt
 aaataggagatcagtcctcctgcagcttgtgggaataaggctttaaatctctccaatttttagctctgtgagatg
 gcactgggggaaacagaaatgaacggactagtgtcacaagctcaggtgggatggacgagatcacttcaaagg
 tctgtaatcccacgtctataatcccagcactttgggaggccaaggcgggaaaatcacttgaggtcaggagtt
 cgagaccatcctggccaacatggcgaagcctgtctctactaaaaatagaaaattagctcagcgtggtggca
 tgctcctgtagtcccagctactcgtgaggctgagacaggagaatcgtttgaacctgggaggcggaggttgca
 gtgagccaatatcacgccattgcactccagcctggctgacagagtgagactccatctcaaaaaaaaaaaaaa
 aaaaagaattttataaaaatcaggaaataatattagtggttattgttgaattttaactttagaatcatagaaaa
 ctctcctctggcatcattatttagacagctcttgtgagtggttagcaccagaccagcttgcatggttattga
 tttttcagagacactttttgagcttattctctggcagaaaggggaactgcttccctccctatctcgtgtctg
 catactagcttgtctttacaagaagcagaagtagtggaatgtttattcttgaaaataagctttttgcttca
 catgatctagaattttttaaaattagaaaaatgtgcttactgcgtgcccttctgaaactaggaaaaatagcct
 tgtgtttaccaattgtgtggttaggagatgggccaaggcatcaggcttttgaaagtagttgcatttagcat
 aatttccattgccccctgccaatttcatatctgtcacatctaatacagtttaaaataaggggcatcctaagca
 tggagatggtccttggatggtccttggagtttctgtattttcagtagtttcttttttttgagcatacaagacat
 ttattgaaaaattcttgggatcaatacttgtgtaaggaaaggaaggaacaaagcatgattgggcagaggc
 agaagatgacatcaacaaaggccctcgttgcagtagttctttttttttttttttttttttttgtgagttgaggtctc
 gctctgtcaccaggctggagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg
 gattctctgtcttcagcctcctgagtagctgggattacaggcatgcaccaccacacctggctaattgttgtat
 ttttagtagagacggggtttcaccatgttgggtggccagactggtcttgaactcctgacctcaggtgatccg
 tctgccttggcctcccaaagtgttgggaattacagacatgaccactgcacccggcctgttgtcagtagttctt
 aaacatagacactaactgtaggctgacagcctagcagcaaggaccagttaaagaaatgagtagaactgaagt
 gtgcttgagtagtctctggcagtcagcaaaaacttaatgggaatcatggtaggccaaatgttctgcagtagttt
 caaaagctgcatgggttttgagaggcttttgggtccatcactcaccttaggttgttctacaagcacatcagcc
 tgccccaatttaaacaggggcagttagtaatggtgtaataaagagctctgcattgttcattccttcaacaaata
 ctggctataatgtttcagcactgtggatgcaaagtgagcaggataaacaggctcttctttcaagcttgtgg
 tccactggaccacgtatgaagtagaatagttttaggtccagaaaggcaattaaagtaaaatagaccaagaaga
 ggctctctagtggttttggtataaagaaaagataagaatgatttagaattggcctatcaatgagataagagg
 cctggcttcttggcactctgctctagggcaagtaaaatggagaattccaaattctgaaattgttagaacata
 gttctgtgtcttagttaaatatctacacttacagataaaatagcataaaatgctttctccccatatttcagccc
 agtccacttaaaagacaacataaaattgcaaaatagttaggtgttctcatctaataaaaagtggttccagga
 attcagactctggattcctgtttgccaatcatgtgtgtccactcttaagaaaacagagttggactctggattt
 ttctttgcaagagggacaagagatgtgggagatctgtgagttaatgcaactgcaggttttaagtgtcctgtca
 ttgtgctcttagtgctttgatacattctgagtttcagttaaagagacctgatgcattggactgttgcaatggaac
 ctgttttaagatcttcaaagctgtattgatatgaagtctccaaaagacttcaaggacccagcttccaatct
 tcataatcctcttgtgcttgtctctcttcttgcagtaaatgcttccag**GTATTTTTGCAAGGCTACCAGTTACA**

TTTGACAAGTCTGTGCAATGGATCAAAATCAGAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCT
 TTGTGGCCTACCAAGGGAGAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCCTGAAGCC
 AATCCTGgtgagtagacttgcctactggagaaaacttcaagcactaatgctttcggaatgtgaggcttttccct
 tggacagcatgactttgtttgtagaaaagtacggctggctgggagttgtgatataattttagttcagtggt
 attctaagtgttcttagtggttctttcagacttttggccatctcccaaaggtgaatgggaagaataagctg
 ggtgtggctgagtttaagccaaaagtttttgtgctgtttcaatcagagaagacctgctttttcatgtttt
 tactattataataactaagcaagagctcatttgaaaacagagttcttcatatttaaaaaaaaaaagtcctgaa
 accattgatgggaagatggatatctattttatgtttaaaaacccatcataaagatgacattgtgggctgtcac
 agttggaaggccctggaattagatgagaccacactattttagcttacttagtaataacattgcaaagaaaaat
 tccgacgaagttttttcagcctaggaatcaatagttcagagaagcactctatgagaatacccattcattctt
 aacaaaaaataactgggtgagcctgagcagtttgggtcatcagagtggtttatatagttccagaacaaatagt
 ctctaggtgttctgagagctctgggtgaaattcctctcgctaccccaaacatcatctttaatatccaggatt
 ctggtttttctactcaccagatagattctcttaaaaccagggaagattcctggaggaaggatgtatctggaa
 agagatgttcttattataataaaaatgaaattgtaatactcttgattttgtgcagcagcaattctttatag
 agagttggctcctcccagagaattaagaatactcagtttctggacctgttcccagatcataccctagaatgt
 gaccttagaaacacacttcaggattcatacctttgattgaccatcaaaaagtttttgtatcgccagggtgtg
 gtggctcacgcctgtaatcccaccacttttcggatgcccagggcgggcagatcacgtgaggtcagcagtttgag
 accagcctggccaacatgggtgaaacctgtctctacaaaaatacaaaaaatagccgggctgatgggtgggca
 cctgtaattccagctactcgggaggtcaggcaggaggtatcgcttgagcctaggaggtggaggtcagtgga
 gctgagatctgtctcactctgttggccaggctggaggtgagtgagtgagtgagtgagtgagtgagtgagtgag
 tctctcatgtcttaagtgtattctcatgcctcagcctccgagtacctgggactacaggcagcctgbcaccacgcc
 cagctattttttgtatttttztagtagacatagggtttcactatgttgcccagctgggtctcgaactcctgagc
 tcaagtgtctgcccacctcgggctcccaaagtgttgggttaaaaggcatgagtcacgtgctgtgtgtgtgtgtgt
 tgttataatttttaagtaaggtatattttctctacagggtatctttgcaaccctaagtaazctggcctaaaaag
 ttagagaagctgacttgtgcagacatttgcagcctgttgggtcttttttgtgctgtgaatcatagaggggtgaa
 aggttattatgaatggtaaaaaactttgttataaaaaccattttcttgactgttttgggctgcttactgca
 tgacaaatgtctcaccctttcagctggaatgattgaaattttggaaaagatgggtgtttttagaagacattgt
 aatttgttccgggt
 aaccgggtaaaaatccctccactcacacaatgcctgaattatactcatagtagaatgactgttttagccctcat
 catctgataaataacagctcagggtttcaacctgacagtatctctctgggaggattagcagcgtgacagagtg
 cagggaaatgcaccttcagaacctcagctacactgt
 atgcgttgggtttatgaccaggtattgattaagggtggctactaccaggtgctttctgcatactcgggtttgt
 ggagcactcaggttctgcttctgccccctctgctgttaccagagacctctcttcaaaatggggctctttagat
 tagagtagaatgagtgatcaggattgttttgtgtaaatgatttctgaggaaggcttttaggtgaaatgac
 ttccaaacattttgaaatgtgactcttacttattgaattaaagcagggccttaattggaatgctgggactgat
 acttgatttgcattaaagcagcctttttctattgt
 gatgtgacatgtgatgacattgcacatgggcaggttaactgtgccaagaagtgcagcagtagcagcaaccgga
 gatgcaaaagcccaacatgatggggagagaaaactcttcttcaatatgtgcttctgtacccaaagtggaaatt
 cagcagagacatatatttgaacattttctcttttgt
 attgtgagtttctcctgggctccttcagaatctgggtgtctgtgaaagcagtggttttggcaacatggggaa
 agtatggcagtggtgggagggtcagctgggtctgggtttgaatatgtcatttgaatatttaccagcattgat
 gtcggataaaattatttagtccctgtaagcctcagttttctcttctctacatacacataatatatttgactc
 tttgttgtgattatttggttacacatatgaagagcctgggtgtggggcctggcacacaatagggtgctcaataaa
 tagaagttgataaatttaattgacatgagtagtagaattatgtccttgaaaacaattgctgcaagatagaag
 ttttcagccaggcacagtggtcacatctgttgtaatcccagcatattgtgggggcccaggcggatgaatca
 cttgaggccaggagtccaagaccagcctggccaacgtgggtgaaatccctctctactaaaaatacacatatt
 tgccaggcaggcgtgggtggcgcacacctgtaatcccagctactgaagaggctgaggcacaagaatcgcttga
 acccaggaggtggaggttgacgtgagctgagatcactccactgcattccagccagcgtgacagagtgagact
 ctgtctcagaaaaagaaaaaaaagatagaagttttctctgtagatcagtggttagaactcataccaagcgaa
 gtggctcctgggtgagttttcagtgaaaaactgcattcttgcctcagatattgtcaagacttttcacccaaaga
 ttcttattttatgtctcagtcctgtaccttgtgtgaaaataactggatgtcagaacgctgttgtgttttta
 aagttccctgggggttaagagcagtttccattattgt
 tgagcaatattttattcagttcttattatgtgtcagggtattttctaggagctggactcaactcaaaagatatc
 ctttgatgagaacaaagggtgggtggatataatgaaatattatctgtgggataaatgcacttagtcatgaggg
 agacttgttatggagtgcgctcattgtattttgtactgttgagtttaacaactctaggaggagctcagggcca
 cctggcaggggcttcttttctgt
 cttttcttaactttctcttgggtccttccctaaagcagcatgtacctttcccagagcagggagagggccacct

agtttcaacatgaaaaaaggtacagagaataacataaagaacactcctggctgggtgtgggtggctcacgcct
gcaatcccagcactttgggagctctgaggcagccagatcactggaggtcaggagttggagaccagcttgcca
acatgggtgaaacactgcctctactgaaaatacaagaattagccaggcatgggtggcgtgcacttgtaatcca
gctacgtgagagactgaggcaggagaattgcttgaaccaggagggggaggttgcagtgcagctgagatcaca
ccactgcactctagcctgggtgacagagtgcagactccgtcttcaaacaaacaaacaaacaaaagaacactc
ctgtaccatcatccatcattttgccgtgctgactccaggttctatttaagaaataaaacattacaggtacag
ctgatgccacctctgtttccctagctcattcttcagagataactctgtcttgcagttggatgttttaatcc
tctatatcaztgtatacttacattctatgtataacaatatttggtagtggcctaaatgtgttcacattgtat
aagtgtgcatattggcctgccacttcatttgggaattatgttcttgagatttatcaatgttgatcatgtgga
atctgggttaatttttggcatagattctatttttataactaaacttttaaaaatccatgcttctagtctttgg
cttattttttcaggttatggatgttttggatgcacagaaaagtaaaattaagtcagcaaaaatatctgg
ataatccaagctttaaacttgatgtagaatttgaatcatgtgtgttttgttaacctgtgatgtcaatccat
gcctgattgtgtaactccaaccaatattccttggaaaatggaaaattgtttatattgactacagattgccaa
tattattagtaaatgctgagcacttaatctcgaataaagaactagtttaaaaatgattctaacaatggcatt
gactgttctaccttattactcatgggtgggttcagccaatgtttctgttggagacaaaacaaaaacagtc
aaattaacaagcagtcacacccaacatacagactactgataagaaggtcatatcataagatatggcattga
atgtgtgtctgctaattgtaaaaatctgatgccacagcaacttaataaggacctatgtttacatccatgct
caattacattcctgggttaaacagtcagctttaggcctgtgtgtgctggagttttgctgaagtgtggg
gcttttaagagaaggagaataagcttgcctcagagtttaagaaatttaaaactaaaagtcctaaagatgttggg
aaaactattgccccttgaagatgtaaatcattaaagttggagaagacctttatacaaaacagacccattc
actgatttgtacccttcaggagacagatgaccggttaatggtagacaatgggtgaatgttgggttgggtttt
tagaaacatctgcacttggtagactactgtatctaattgggtgtgacaaacctggcaccccatgtgttggcac
catcttgggtcctactcagggccaggtgaaccgagtggtccttctactgcttcagagctgaagcagattgt
agtatgcccagcagacacagaatataccaccaagcagcttggcgcaagcactactgggaagggaggtttgt
gaacatgggtgctggttctcaaacctcagtgctcaaaagagctcctcaggattcctggatcacactcttaa
cttctcattcagtaggtgttagctggttgagaatctgcatttttttttttttctgagaccagctcact
ctgatgccaggtggagtgatttggcgccatcttggctcactgcaacctctgcctcccaggttcaagcaat
tttctgcctcagcctctctagtagctgggattacaagcacatgccaccatgcctggctaatttttgtactt
ttagtagagaggggttttgcgcgtgttggccaggtcgtccttgaactcctgacctcaggtgatcccccact
ttggcttcccaaagtgttgagattacaggtgtgagccaccatgccagcctgaatctgcataatttaacaagc
accacaggtgattctgatacagtagctcccaaacctcacagtggttagtgaatcccagtcattttacaattct
gccatgattttgggtcatattcaagtgcagctggtagcatttttagttaatatatttttaaaatagtcact
tcttttgggataattaaatttaattacaaggaagctataccactgctgtaaaaacatcacctgctttaaga
gaaggtacataatgaatatacattaaagataaagatgtatatgtgtgtgtgtgtgcacataaagtatacaca
tacctaccatagggattgagttttccttcaggtttttcaactgaaatgtcaactttgaggccagtttaatatg
tgtaagatatatgtgtgtatgtatgtctatacatatagacataacgtaaaaacatacatggatgcataatg
tatactatacacaacctattatgcataatcatgtatatattccacttagtattatcttztattttgccg
tttggcaaatgctcagtaaaaagaaaggggttagaaggggagaaaggcattttatcccaagccttcaggaatc
aggatgaggatgtcttcaccttctggttggggagtagtatacaattagagacagcacatttggagztgtggc
tgatatgctgtgtgatgatagcttagctctctgcctagcagaggaaggacatttcaatagaagaaaaagt
taagaccttgcgagaaaacagagaaaaggatgtttgtctttttaagaagttgaaaacctgttgcagacaaa
agcctccagttttggcagtaaaactttcatgcaaggaagaaaaaggcaggggatgacattgttgacaattg
tgaggaattaccatgtgccaggcactgtgcgaggggctttagtataatctctagtttttagtgcttataaaa
actctgtgatattgtgcacagcattttaaactttgctgcatagtcgagaaaatggaaggatggggaatttgag
tcatttggccagggttctatagctacccaggttcccatgactggagaattggggcacaggggtggcggggga
gagtgagtgacaagaatcctaacaatcttatttccattgagtccttataaaaagaagtggaattaactaccag
tttttaagtttttctaaatttaggttatgtggatctggcggttcttgttttgcctgggttgggtttgttt
ttgctatgctgtcttgaacatctgtcatctttagggcctaaccggtaaacacaaaaacactttacctctata
gctttcaattaaagatctctcagtttgtgtttgtaatatgtttccaggcaagttctccctaggttcggcttct
agtgtgttaaccttttagttataaaagtgaacccaaagagagaaagtagaaaacaaacacctcacctgttttg
ctcatgaattactctctatggaaggacaatcatgaacacctctgcgtatcacagaggcctatctgagctg
acgtttaaggagaccgcgtaggtccctttgaggactgtgaatgtgggagtcctgggactcgtggaagaac
ccgttcagaagagatgaatgagctggacaagttcttctatagaaccttttaggcaggttttcttagaaatgc
acattagagattatgcttggatattgtgatgcagaatgataactcaatcccttctgcatttggaaattctct
ttgaagaaaaacatcccaggcagctatttctcagagatagtgagtcacagccacttctagacattttcttgt
gtagctacattataatttccagcagctctctgatagcaaatgtcaaaatagcccaaccttctctaaact
tcagagatgtctgatattgaataaaacaatgctcatagaacatcaagaaaggtggattttccctgg

atacttttttctgtgcttgacaaataacagtgaaagaaactgatctcacgtctttttctcttttgaagcctgaa
 cactcagaacccaacttgaggctcctcagctatagcaattctgacttcacagctctgtaaattattgttcttt
 tttttcttttagcttatgctttctgcccataatttatcttttccctgttctaataatgaattattgtcctatatctg
 ctgtgcagtttaggtgacatataacagcaattaaatatatgaattggtacatataaagatttgactaaaactc
 gatgtaaaaataagtggtctacattcaatttccagtggttagaaacagtgctgacttgaaacagagtgacagaa
 ttccatctttccctatttttgacagcttttaaactttatattttcttcttcttctgtgagccgtcattaactt
 tgtttctcaaagzccattcccgtattacccatcttgcagacgcagacagatttggaatttgcggtcagagt
 tgtattggacacatccccccagcccacatgagatccttttaattctattgcatattaactagttttaagtaca
 atattcctacttcattttaaaccattaatcaaagaatgagtttgaaaatgaacaaaatgcaaacttacagtt
 agaaataattgtagtgtcttttagtttgggttaggagtcggtttctgtttgtttaaactcaagattgtgaaca
 gttttaattcacttggtttatttccaatagagatttcaggtttacatttgaattcagaaacaaagttttcttt
 ctcatcag**AGAACACTAAACTCTACATCTCCCTTCCCGAGCAAGGAGCTGGCTGAAGCCACAAAAACATT**
GCTGCATAGTCTTGGGACTCTGGCCAGGAGgtaagttgtgtctttccagtagcaggaagcggatcatccac
 tgtatcagtattttcttctgagtcgtggcaagaggtccttttgagttgaatatcacatgggatgtaatatc
 aattttcaaagtataagtgtgtaaaacaataatgttttgatttcttattttagaatgaagaaacctaaaa
 ctcatagatgtctcagagctaatttggttagtggttaacagctggatatctagtttagaaccttctccatttt
 ttctttttgcccctaggtaatcatatcttgaagaggagaattatctctgccactgcccagcactgctt
 ttgtctgaccagcaatttctccatattgcttcttcagtagcaaggccaatcattttaccaacacacatgctt
 gctaactaacaggaataacgtggttacccttaattcagccctttcccttgaaagcatctggcttctgaggttc
 aactatgggaatatggtctcttaataacacatttaagttgagttgtgcttttaggtccacatggtgacaaatgt
 atcagagtaactctctgtccttaggatcagagggcctgtgagcacttgcaaaagcagtttagctctgactcccag
 ccagtgacacactccacctttctgactcccagccttgtctcaaattaggcttggaagcaggaactgtctggt
 gtccccagcataggaagctgagccagggggcagtgctcacaaacaatacagactttaacgtgtaggatatt
 ggaaaataataatttgtggggaaattgtctcagacttggtccacccttatttttagctgcttctctaaccg
 tttttcttttttgggtgcttcttaacctacccttttttgggtgcttgcattttttcaaataatcaaa
 aacgaactttatgttttctaacaatgaaagtattgcatgttcattgtggaaaatgctgaagacttggaat
 acaaaaatgctgagatcaaacactattgatacgttagtgattttcttctgtcctgtctactttcttctt
 tgaattctgtctcacgtgtttctgactgatgaggtctgacttttgggttccctttccagaggagaagccttct
 ttccagcttgccatttgttaccctgggttatgaaggctggtaaccttttttactaggttagagaagctggacca
 ctgggggttcttccagggggagaatgagaaagagaaactgttttgcaagtccttagctatttctctaggccc
 tgttagctgacattgacatgccttgcatgtctgcagatcccctcgcagccctctgtcccttgttcatttc
 tggccttagagaaagcaaacaggggtctgtaacaggggaggtgcctctaaactcaggggttggttacagct
 gttttcacttacatcactggccctgggttttttttttttctggcattaaaaaaaaaattggaagcaggtg
 atgttcccatgtctgatgtggtggaaactctccaagtgaacaatatagctttttcttggcagctgtttcttg
 tgcctgtctgtcctgggtccaggacaagcaaggaccatctgcctctttcaatagaacacctccagatccct
 ttgatcaaaagttactcattgtctgacttgctatttctgtgagataaaatgggagaagatcaataaatgcact
 tgtttgtccagtcagczgtgtggaaagtgtataattttgaccaaagcacaacccttgaaaggaaaagaaaa
 gggagtgaatgtcttctgagaagctgcctaggctcagacagtggtcacccatttccctgtatgtccacatga
 caaacctgagtgggtctcatctgctcattttgcagatggcaccagggtcagaaaggttaggcaacttttc
 cagtcaccaatgagtttaattgacaaaactgggattcaaacccagaactgttgattccaaagcctgtgttg
 ttgcctgtctgtgaaaaactccagtagcagactggaatagaaaggagaaccttccaagaaagaaaatacgca
 ctagcagaacctggaaattgggagggaatgaggacttgaggaataagatgaatgaaagctgacctgagtttc
 acatctgggtgatgggaaggaggagcagggaggcagcatctcagatgtccaccagcaccgaccagctgcct
 ggcattgtcaggtgttgaggactcagcagtgaaacagctaaacttctctgctttcttggggcacgtataggg
 gagagacagaaacaaacaggtcagtgtaaatgccacaggaggatataatgcagtgaagaaaaagcagggt
 aggggcatagagcatgagaaggtgctttttttaaaggggktgattaggaaagctctctctaaggtgacagtt
 ggacctgaaggagatgatagcatgtctgtggtgaggggaaggaaactccgaacaggaagaatggcagatacaa
 agacattgatgtagagcatgcctaaggaatgtgtttaaggaccagggaaagtgaagcaggtgggggggag
 gagaggagctcagagcaggaggaggtgagtgccatacaggcctggcaagactttggattcctgtctgggtgag
 atgagaatccagcggagggttgagggaggggacatgatgtgatctagagtttagactgtttacactctggt
 tgttgggttgagaagagactgggatgggggaaaggaggacaaaggacattgtgctggattgagaagcag
 aagtcagtttcattcattcactcaaccgatgatgttcaaataccaccatcatccgtgggctaaaggatgaag
 agccatccctccctgagagtcaggaagcacttcccagataaagtttgagtgtagctgaggtgtaggagaa
 agagtaagagtttaccctgaaacgggtgctgggaagagtcataagtttggaaataactcaataatttatggt
 gcttcttttagaaagatttgctggctttatgtgggaagaaatttkttttttgattggggagtggtgggttg
 tggtaggctgcctgtggaaagagaagtgagttttgactcactgttatttaaaaatctctagggtgttc
 caataagcaacaaaaggcaaatggcctggttctctgtccctttctgtctgtatgcctcgtacaggttatg

[illegible]

gatttttttataaagcagcagttttcagacccttgagactccaattcggtagaaccagagcttcatcttct
ctgtcgaagctgtgacaggagttgcaaagcctctccttttctgagtttgagctgctgttttccggca
gcacatctgtgcaggcctctgctcggccctctggzatctgctgattgagcagcggattgatctgtccttc
tctttcgtgttgacccatgtgaggaaccaactggcaagggaagaaatggaaataggcctcctttgcat
catgacctgtacatcctgcaattggaaaagattgtacttttagttggtttaaccagcagcattatctttaa
actaagcagtaagaaggaattaggttttatgtgggatcaacagactgggtctcaaaagaggaaggtgataga
acacagtgaggaggaggagggtgcactagaaacagaggcctatgctttcattctggctttgctacttaata
ctgtgtgacccaatcttagagacttaacctctctgaacttccattttctcatgtataaaatgggaaatatta
aaggatactcactgggctgggtggcttggtgcctgtaatcccagcacttggggaggttgaggtgggaggatcac
ttgagcccaggtgttcaagaccagcccagggaacatggcaagactctgtctctatgaaaaatataaaatta
gccaggtgtgggtgggtgtgcacctgtagtcttagctacttggtaggtgagatgggaggatcacttgggcttg
ggaggtcaaggctgagggtgagctgtgattccatcactgcactccagcccggcgagagcagacactgaa
tccaaacgacaacaacaacaaaggcaaaaaataaaagtgcctcctttatggagttgtgtaagggtgaagca
tatacactattcaacatagtaactatataaaggaagtattgttgttactgtagttaataaccattaagt
agatgtttcgtatagtggaaagcacatggactctgaattcagactgggtctgactttgagctctcagctccaca
tctagtaataactatgaccaagcctgggttaaaatcatgttttttttcttcagcctcagctcttctcacatat
aaaatagggacactgtcatttacctcagttttctgtgaggataaaacaacgacagtgtatatgcaagtattt
tgtaaaatctttagtgctcctcagagtttagttgggtgttactacttgtactttctcactggaatggcagAT
GCTGTTGGACAGCAGGGACAATGACCACCTTTGGGAACAGCAGTTGGATGGCTTAGATTGGACAGCCCAAGA
CATCGTGGGCGTTTTTGGCCAAGCACCCAGAGGATGTCCAGTCCAGTAATGGTTCTGTGTACACCTGGAGAGA
AGCTTTTCAACGAGACTAACCAGGCAATCCGGACCATATCTCGCTTCATGGAGgtgaatctgttgcctgggatc
atthagaaaagacttaacggcttcttctctgagacgttacaataagggttcaggcaggaggcaagtttagaa
ataatgtatagtctcatttacaaaactatccctcaagcctaacacaggatttgataacaaaaggcacttaata
aaatgttagttgagtggttgatgagtaataaaactctagcttttagtaaaattaactctagcttattctatat
aggctcaagagaatatttctacccattttcttcttaggttttctctatctcagtgactaatggtagcaagcat
tcccttaaaaaggcattatttgtgaaacttatctaaaatcgaattcgggtccaattaaattttgaatattt
atattaaaatattatattagtagggatgggtgaagaggtgttttgggtctgggtgggttagttgctatgact
cagaattgctaagaaaacagaaaagtaagataagatcattgttttaacctcttttctccacaaaatcaata
aataacatatccctaaattactcttagaatttctcttaaattgcagtgaaaaacaaaatccttcattcttg
gttgaagggttggaactacgttagagaggattagagagagaggatgagcaatcgtgtagtcagcccttggc
tcttagtgtaggatttgtctcagccactgcttgttgcctggctgccaacgttctcatgaaggctgttcttc
tatcagTGTGTCAACCTGAACAAGCTAGAACCATAGCAACAGAAGTCTGGCTCATCAACAAGTCCATGGAG
CTGCTGGATGAGAGGAAGTTCTGGGCTGGTATTGTGTTCACTGGAATTACTCCAGGCAGCATTGAGCTGCCC
CATCATGTCAAGTACAAGATCCGAATGGACATTGACAATTGAGGAGGACAAATAAAATCAAGGATGGgttaa
gtggaatcccatcacaccagcctgggtcttggggaggtccagagcacctattatattaggacaagaggtactt
tattttaactaaaaatttggtagaaatttcaacaacaacaaaaaactcaacttgggtgtcatgattttgggtg
aaattggtacatgacttgcctggaaggttttcataggtcataaaataacagtatcttttgatttagcatttc
tactcaagggaatttaattccaggaatttgggtggcaggcacctgtaatcccagctactcgggaggctgaggc
aggagaattgcttgaacccaggaggcagaggttgagtgagctaagatcgcattcactccgcctggg
caataagagtgaactccatctcaaaaaaaaaaagatacaaaaatagaaaaaggggcttggtaagggtag
tagggtttgggcaatt
tggtttgggttttagGTACTGGGACCCTGGTCTCGAGCTGACCCCTTTGAGGACATGCGGTACGTCTGGGGG
GGCTTCGCCTACTTGCAGGATGTGGTGGAGCAGGCAATCATCAGGGTGCTGACGGGCACCGAGAAGAAACT
GGTGTCTATATGCAACAGATGCCCTATCCCTGTTACGTTGATGACATgtaagttacctgcaagccactgtt
ttaaccagtttatactgtgccagatgggggtgtatatatgtgtgtgcatgtgcatgcatgtgtgaatgatct
ggaaataagatgccagatgtaagttgtcaacagttgcagccacatgacagacatagatatatgtgcacacac
tagtaaacctctttccttctcatccatgggttgcacttttatcttttttttttttttttttttttttttttt
gagctctcgtctgacgcccaggctggagtgcagtggctcgatctcggtcactgcaacctttgcctccggg
ttcaagctattctcctgctcagctccacagtagctgggactacaggctcatgctgccacgcccggctgac
tttttgatttttagtagagacgaggttcccatgttaccaggttagacttcaactcctgagctcaggcaa
tccacctccttggcctcccaaagtgtgggattacaggtgtgagccactgcacccagccaccactttaat
tttttacactctacccttttgggtcaaaatttgcctcaactctgcaagcttaaaatgtgtcatgacaaacacatg
caagcacatactcacacatagatgcagaaacagcgtctaaacttataaaagcacagtttatgtaaatgtgtg
cacttcttctccttaggtggtaaacacatttcaaaacaccccaataaaaactgaacaaagcttcttctct
tagactttttagaaaatctttcagtgctgagtcactaagctgccaagttctcattgtgggaactatgccttt
ggatgtaatgatttcttctaagacaatgggcggaggtgtagttattgcagacatctgaaatatgtaatgttt
cttcagattctggaattctcttattctctgtggttggtgggtgggtgggtgggtgggtgggtgggtgggt

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TTTGGCTGTGAGTACTTTGCCCTTTTGGAGGAGCAGGGCATTGGAGTGCAGTGGGACAACCTGTTTGAGAGT
CCTGTGGAGGAAGATGGCTTCAATCTCACCACCTTCGGTCTCCATGATGCTGTTTGACACCTTCCTCTATGGG
GTGATGACCTGGTACATTGAGGCTGTCTTTCCAGgtacactgctttgggcatctggttgaaaatatgactt
ctagctgatgtcctttcttgtgctagaatctctgcagtgcatgggcttccctgggaagtgggttgggctat
agatctatagtaaacagatagtcacaggacaggcagctgatgctgaaagtacaattgtcactacttgtacag
cacttgcttcttgaaaactgtgtgccaggcagcatzgc aaaatgttttatacacattgcttcatttaattct
cacaaggcztactzctgaagtagttactataataaccagcaattttcaaagagagaactgtgactcaaaga
cgtaagtaaccagctttgggtcacacaactgttaaatgttggtacgtggaggtgaatccacttcgggtacac
tgggtcaataagcccaggcgaatcctcccaatgctcacccaattctgtatttctgtgtcctcagaggggta
caactaggagaggttctgtttcctgagtacaggtgttaataataataataactagctcctaaggcctgctg
tgatttaattagcattcaataaaaaattcatgttgaaattttctttagtacttctttcttaataataacatc
ttcttgaccaagtccaagaggaacctgcgttggacagttttcatatgagatcaaattctgagagagcaagat
ttaaccctttttgggtcaccttctgatcctcccctaaggaggtatacatgaaatatttattactcctgctg
aacttctttcattgaatatgcaattttgcagcatgcagattctggatttaattctgagtccttaacttactg
gctgagggaccttggataggctccttatccctcagtttctcatctctaaaatggggatggcacctgccccg
tgggttgttggaaggacttacagaggtgcagaatgtacgttgatcatagcaggttccagcaaatgttagctc
cctctttcccccacatccattcaaactgttctcttccaaaggatgtgtcaaggaggaaatggacctggctg
ggaaaccctcagaatactgggatgatgctgagctgtggtcctacacctgtgctttgcttccagGCCAGTACGGA
ATTCCCAGGCCCTGGTATTTTCTTGCACCAAGTCTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCAC
CCTGGTTCCAACCAAGAGAATATCAGAAAgaagtgtgctgttgacctcctgctctttctttaacctagtgc
tctgctcctctgtcactgttgggggcaagcagatgtctcctgctttctaaaagactgtgaaaccactccagg
ggcagagaaatcacatgcagtgctcctttccaaactcctccatgccatttatgtccaatgctgttgacctat
tgggagttcacggtctcgatccctgagggacattttcttgtgtcttggcttctagaagagtatcttttac
ttgccccctcccaaacacacatttcatggtctcctaacaagctagaagaaagaggtaaagacaagcgtgatt
gtggaacctatagcctcgtgctgctgctgtgacatgggtgacctgtgtatcagcctgtgtgggctgagaccaag
tggctaccacagagctcagcctatgcttcataatgtaatcattaccagatccctaactcctctcttggctct
taactgcagacagagatgtccacagctcatcaaaggctctgccttctgggttcttctgtgcttagagtggctt
cctaataatttaatagggtcccttttctgcccagctctcttctgtgcccacccctgattgccccttggtaaaagt
atgatgccccttagtgtagcacgcttgcctgctgttcctaatactcttctcctacctcctctttacacctag
ctcctgtttcagtcacctagaaatgctcacagtcgctggaatatgtcatgttcttccacacctccatgcctt
tgtaggtactgtttgctctcacaggagaactttctctctaacttgccctatcttctcaactcctcctttctct
ccaagatctagttccggatccccctccccctgagcatccctccttgggttctcaggtagtcagtcactctctgcc
ctgaacttccatggcacgtgaaagaaaaatctttttattttaaaacaattacagactcacaagaagtaataca
aattacatgagggggttcccttaaaccctttcatccagtttccccaatggtagcagcatgtgtaactgtagaa
tagtatcaaaaccatgaaattgacataggtacaattcacaaaccttcttcagatttccactagctttatgtgc
gctcatttgtgtgtgtgtgtgcgtatttagttctatgcaattttatcatgtgtgaattcatgtaattactag
ctcagtcagctgcagaaatatctcattgtcacaaagctccttcatgctaccccttaaztgccacazagccz
acctcccttcttctcctcagttcctgacacctgtcaaccactaatgcgttccctgctttttacagttttattatt
tctagaattgttacataaatggaaccatacagtaggtatccctttgatactggcttttttttttttttctactc
agcagtatcccttagatctatccaagttgtgtgtgtcaacagttcattcctcttctactgctgagtagtgtt
ccctgggaggggtgtatcacagttccatggcatttttagatgtatttttaaacagctttcagcatcctcta
ttttaattgttcatcaagtcctttttcccaatagactctgaatgctcctttatcatcgtattcccacacca
acatcagtagcccaaataggccctaaataaacatttatagcctcctgctgctgagaaaccagggtggacat
ggagagaaggcacttctgaaagttcaagcgcagtgcsctgtgtccttacactccactcctcagtgctttctg
tgggttcatttctgtcttctctcctgtcacagTCTGCATGGAGGAGGAACCCACCCACTTGAAGCTGGGCGT
GTCCATTGAGAACCCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTGTGATGGCCTGGCACTGAATTT
TTATGAGGGCCAGATCACCTCCTTCTGGGCCACAATGGAGCGGGGAAGACGACCACCATgtagaagaggg
tgtggttcccgcagaatcagccacaggaggggtctgcagtagaggttagaaatttataccttaggaaacctg
ctgatccctgggccaagggaaggagcacatgaggagttgccgaatgtgaacatgttatctaatcatgagtg
ctttccacgtgctagtttgcagatgttatttcttcagcctaaaacaagctggggcctcagatgacctttcc
catgtagttcacagaattctgcagtggtcttggaaacctgcagccagaaaagatagattacatatgttggag
ggagttggtaattcccaggaactctgtcttaagcagatgtgagaagcacctgtgagacgcaatcaagctgg
gcagctggcttgattgccttccctgacacctcaaggaccttacagtggttagtatcaggaggggtcaggggc
tgtaaagcaccagcgttagcctcagtggtcttcagcacgattcctcaaccattctaaccattccaaagggtta
tatctttgggggggtgacattcttttctgttttcttttaattcttttttaaaacatagaattaataatta
tgagcttttcagaagatttttaaaaggcagtcagaaatcctactacctaacaacaaaaattgtttttatcttt
gaataatatgttcttgtttgtccattttccatgcatgcgaggttaggcatacaaaatacattttttaagaa

tacttttcattgcaaattggaacttcgtttaaaaaatgctcataactaaaattggcattttctaaccocataggc
ccactttagttatattaccgaagcaaaaggacagctttgctttgtgtgggtctggtaggggttcattagaaaag
gaatggggggcgggtgggaggggttggtgttctgttctctctgcagactgaatggagcatctagaggttaagggta
gggtcaaccctgacttctgtacttctaatttttgcctcag**GTCAATCCTGACCGGGTTGTTCCCCCGACC**
TCGGGCACCGCCTACATCCTGGGAAAAGACATTGCTCTGAGATGAGCACCATCCGGCAGAACCTGGGGGTC
TGTCCCCAGCATAACGTGCTGTTTGACATgtgagtaccagcagcacgttaagaataggccttttctggatgt
gtgtgtgtcatgccatcatgggaggagtgggacttaagcattttactttgctgtgtttttgttttttctttt
tttctttttttatttttttgagatggagtctcgtctgttagccaggtggactgtagtggcgcgatctcggct
cactgcaaccttggcctcccaggttcaagcgattctcctgcctcagcctcccagtagctgggactctaggg
acacaccaccatgcccagctaatttttgtgttttttagtagagacgggtttcaccatgttggccaggatggg
ctcaatgtcttgacctcgtgatccgcccacctcggctctcccaaagtgtgggaacacaggcatgagccactg
tgtctggccacattttacttttctttgaatatggcaggtcacctccgtgaacaccttgagacctagtgtgtc
tttgatttttaggagaagtgggaggtgaatggttagctgtagaggtgacatcagcccagccagtggtggg
gcttgggaaacattgcttccattattgtcatgtgtgagggccctttagcccatcctctccccccgcccacc
tccttattgaggcctggagcagacttcccagacctggtagtgttcagggccctggtatgatggacctatat
ttgctgtttaagacatttgcctccactcaggttgtcccatcagccataaggcccccaggagcccggtgtgat
ggagcagagagagacctgagctctgcaatcttgggcaaggcttttcccttatgtttcttctatctaagt
aacagctggggctcatgtgtctcctcctcatctaagtgaacacatggggctcatgtgcagggctcctccccg
ctttcagagcctgaggtcccctgaggtcaggaaggctgtccaggtgagtgcagagctgacttcttgggtg
acgtgctgtgggacagcccattaaagaccacatcttggggccctgaaattgaaagtgttaactgcctgggtg
catggtggccacggcctgtggaacaggttgggaagcgtatgtcacctttcactttgatttctgagcagct
catgtggttgcctcactgttgttctaccttgaatcttgaagattattttccagaaattgataaagtattttta
aaaagcacggggagagaaaaatatgccattctcatctgttctgggcccaggggacactgtattctgggggtat
ccagtaggggcccagagctgacctgcctcccgtgtcccag**GCTGACTGTGAAGAACACATCTGGTTCTATGC**
CCGCTTGAAAGGGCTCTCTGAGAAGCACGTGAAGGCGGAGATGGAGCAGATGGCCCTGGATGTTGGTTTGCC
ATCAAGCAAGCTGAAAAGCAAAACAAGCCAGCTGTCTAGgtgcggcccagagctaccttccctatccctctcc
cctcctcctccggtacacacatgcccaggaaaatcagcactgcccaggggtcccaggtgggtgcggttgg
taacagaaaactgtccctggctgtgcccctaggtcctctgccttactcactgtctggggctggtcctggag
tttgtcttgcctgtgtttttttag**GTGGAATGCAGAGAAAGCTATCTGTGGCCTTGGCCTTGTCTGGGGGA**
TCTAAGGTTGTCAATTCTGGATGAACCCACAGCTGGTGTGGACCCTTACTCCCGCAGGGGAATATGGGAGCTG
CTGCTGAAATACCGACAAGgtgcctgatgtgtattttattctgagtaaatggactgagagagagcggggggct
tttgagaagtgtggctgtatctcatggctaggttctgtgaagccatgggatactcttctgttakcacagaa
gagataaagggcattgagactgagattcctgagaggagatgctgtgtctttattcatcttttgcctccaac
atggtgcactaaatttatggttagttgaaaggggtggatgcttaaatgaatggaagcggagaggggaggaag
acgattgggctctctggttagagatctgatgtggtacagtatgaggagcacaggcaggttggagccaactc
tggcztgccctgagacattgggaaagtcaaacctgcctcacttcttggcgataataatagtgtgczt
tacctcatagaggattaaattaaatgagaatgcacacaaaccacctaagcacaatgcctggcatatagcaagt
tcccaataaaaatgcztactgttcttacctctgtgaggatgtggtacctatatatacaaaagccttggcattc
taggrrtcataagccatacaggggtgaaaggtggctccaggtctcttccagtgcttaccttgcctgctaatactc
tctagtccctgtcactgtgacaaatcagaactgagaggcctcacctgtcccacatccttgtgtttgtgctg
gcag**GCCGCACCATTAATTCTCTCTACACACCATGGATGAAGCGGACGTCCTGGGGGACAGGATTGCCATC**
ATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCCTGTTTCTGAAGAACCAGCTGGGAACAGGCTACTAC
CTGACCTTGGTCAAGAAAGATGTGGAATCCTCCCTCAGTTCTCTGCAGAAACAGTAGTAGCACTGTGTCTATAC
CTGAAAAAGgtgagctgcagctcttgggtgctgggctggtgttgggtctgggcagccaggacttgctggctgtg
aatgatttctccatctccacccttttggcatgttgaaaccacatctccctgctctgttgccctttgaaa
tcatatcatacttaaggcatggaaagctaaggggcccctctgctccattgtgctagttctgttgaatcccgt
tttcttttcttatgaggcacagagagtgtggagaaggtccttagaggacattattatgtcaaagaaaaga
gacttgtcaagaggttaagagccttggctazcaaatgacctgggtzgttccctgctcattacttttcaatctcat
tgaccttaacttttaactataaaaacagccaatatttattaggcactgatttcatgccagagacactctggg
caztgaaagaaagtaatgataatagttaattttatatagcgttgttaccatttacaacctttttttttttt
taacctctatcatctcaattaaagtgcagagagaccctgggaagaaggttaactatatttattatccagatg
agggaagtgaggctttagggaattggtagctgattcaaggtcaccagcaggttaataaacagtggtgggac
cagaccaattaccaggtatgttttctctgtaccgcagtagatgcctgagatttatttgtgtgtgaagcc
agtggtagcctaattgatttacatcccaacctgaaactcctatccacttatttaccttttaagtgcctctta
actcaagtgcagctgaggaccagcagcatcaggatcacttgggaacttgttagaaattcagcaacctgggc
ccagctcagacctaccgaatcagaatctgtgcattttaacaaggttcttgagtgggtgaacacacattaaag
catgagaagcattgaactagacatgtagccaggtaaaaggccttgccctgagatgggtggcaaggccctcattg

cagcattcattggcagggccacagttcttttggcagctctgcttccctgacctttccaccctcaggaagcgaggc
tggtcacacggcacacacatgcccagacaggtcctctgaagccacggctgccagtgcatgtgtcccagggaa
agctttttcccttttagttctcacacaacagagcttcttgggaagccctccccggcggaaggtgctgggtgctctg
ccttgctccgctccctgacccgttctcacctccttcttggccatcag**GAGGACAGTGTCTCTCAGAGCAGTTC**
TGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGCTGACCATCGgtaaggactctggggtttcttattc
aggtggtgctgagcttccccagctgggcagagtggaggcagaggagaggtgcagagggctgggtggcgc
tgactcaaggtttgctgctgggctggggctgggtggctgctgggkgtgaggagcagcttgggtggcgggttggcc
taatgcttgctgggggtgctggggctcggtttgggagctagcagggcagtgctccagagagctgagatgatt
gggggttggggaatcccttaggggagtggaactgaataaccagggatgaggagctgagggccaagccaggag
gggtgggatttgagcttagtacataagaagagtgagagcccaggagatgaggaacagccttccagatttttct
tgggtagcgtgtgtaggagggcagtgctaccagtagcatatgtggaacagaagtcttgacccttgctatctc
tgcttagtcctaatggctggcttttccaggaaggcttctgcttccatggactgttagattaaccccttatt
taggtaaatgaggggaacctactttataagcataggaaagggtgaagaatcttttaagattcctttactcaag
ttttcttttgaagaatcccagagcttaggcaatagacaccagactttgagcctcagttatccattcaccat
ccaccacccacccacccatccttccatcctccatcctccattcaccatccaccatccagctgtccac
ccattctacactgagtacctataatgtgctggcttgggtgatacaaaaggatgaataagacatagtcctttcc
tttggccccaacccctcagaccagagatgaacatgtggaatgacctaaacacctggaacagggtgtgggtgatg
agcggcagggcctctgatgagaggggtgggggatggccagcctcactccgaagccctctgagttgattgagc
catctttgcattctgggtccctgag**ATGTCTCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCC**
CGGCTGGTGGAAGACATAGGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGGGAGCCTTT
GTGGAACCTCTTTATGATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTAGTTATGGCATCTCAGAGACG
ACCCTGGAAGAAgtaagtttaagtggtgactgtcggaatatatagcaaggccaaatgtcctaaggccagacc
agtagcctgcatgtgggagcaggattatcatggagtttagtcattgagtttttaggtcatcgacatctgattaa
tgttggtccccagtgagccatttaagatggtagtgggagatagcaggaaagaagtgttttccctctgtaccaca
gtacatgcctgagatttgtgtgttgaaccagtggtacctaacacatttacatcccaaccttaaaactcctat
gcacttatttaccctttaatgagcctctttacttaagtacagtgkgagggaacagcggcatcaggatcacttg
ggaacttggttagaaattcagcaacttggggccagctcagacctaactgaatcagaatcaggagcaattctctg
gtgtgactgtgtcacagccaggtatcaactggattctcatacataggaaatgacaaacgtttatggatggat
agtctacttggtccaggtgctgagatttgttttttgttttttgatttttttttaactactgtgacctcattt
aattctcaaaaaaagatgaaaaaatgaacactcaggaatgctgacatgagattcagaatcaggggttgggg
cttcaaagctccatcctctctttatccatgtaatgcctcccttagagatacaacatcacagacctgaaggc
tgaaggggatataaaagctgtctggccaagtggtctccaagcttgacagtgcagcagaatcacctggggata
ttattaaaaataaacataactaaggttgggttcagggcctgtgaatcagaatttctggaggtgaggccttga
agtctgtatttctattgcatacttggacacagtggtctatagactagagtttggaaatgattgcgctcatt
cagattctctctgatgtttgaattgctgccatcataatttctagtgtcttatttccctcctgctcattctgtc
ttggataacttatcatagtactagcctactcaaagatttagagccacagtcctgaaagaagccacttgactc
attccctgtaggttcagaataaatttctctgcgcagtgctgtcatagcttttttttaatttttttttatt
tttgatgagactggagtttctcttatttccccaaagctggagtgagtggtgagatttgggtcactgcaac
ctccacctcccaggttcaagcgattctcctgcctcagcctcccaagtagctgagattacaagcatgtgctac
cagcgcagctaaattttgtatttttagtagagatgggttttatccatgttgggtcaggctgggtctcagactcc
agacctcaggtgatctgcccgcctcggcctcccaagtgctgggattataggcctgagccacagcgtcagc
cataactttaatttgaataatgattgtctagcttgatagctctcaccactgaggaaatgttctctggcaaaaa
cggcttctctcccaggttaactctgagaaagtgttattaagaaatgtggcttctactttctctgtcttacggg
gctaacatgccactcagtaataataaatcgtggcagtggtgactactctcgtaatgttgggtgcttataatg
ttctcatctctctcattttccag**ATATTCCTCAAGGTGGCCGAAGAGAGTGGGGTGGATGCTGAGACCTCAG**
gtaactgccttgaggagaaatggcacacttaagatagtgccttctgctggcttctcagtgccagagtattg
ttcctttccctttgaattgttctattgcattctcattttagagagttaggttgggtgcagatggggaagggt
tgttttgttgtaataaaataaaagtatttcttctgtgccttcag**ATGGTACCTTGCCAGCAAGAC**
GAAACAGGCGGGCCTTCGGGGACAAGCAGAGCTGTCTTCGCCCCTTCACTGAAGATGATGCTGCTGATCAA
ATGATTCTGACATAGACCCAGgtctgttagggcaagatcaaacagtgctcctactgtttgaatgtgaaattct
ctctcatgctctcacctgttttcttggatggcctttagccaaggtgatagatccctacagagtcctcaagag
aagtgaggaaatggtaaaagccactgttcttgcagcatcgtgcatgtgatcaaacctgaaagagcctatc
catatcacttcttttaagacataaagatgggtgcctcaatcctctgaacctgtatttattatcttttctg
cggggcctagtttcttgtatacattaggtgtttaaattgttgaacaaatattcattcgagtagatgagtgat
tttgaagagctcagaaaggggaatttgcgttagagttaattgtaccctaagacttagatatttgaggctgg
gcatgggtggctcatgccagtaatcccagcgtttagagaggtgaggtgggttagatcacctgaggtcaggagt
ttgagaccagctctgaccaacaaggtgaaacccgtctctactaaatacaaaaaatttagccgagtggtggc

acatgcctgtcatcccagctacttgggaggctgagggcaggagaatcgcttgaacccaggaggcagaggttgc
 agtcagccacggttgcgccattgcactccagactgggcaacaagagtgaactccatctcaaaaaagaaaa
 aaaaagaattagatatatttggatgagtggtcttctgtgtgttaactgagatggagaggagagctaagacat
 caaacaatatattgttaagatgtaaaagcacatcagtttaggtatcattagtttaggacaaggatttctagaaa
 attttaggaacagaaaactttccagttctctcacccctgctcaaagagtgtatggctcttacattatataat
 aactgcctgacttcatacagtatcagtaacttagatcatttgaaatgtgtccacgttttaccaaaatataata
 ggggtgagaagctgagatgctaattgccattgtgtattctcaaatatgtcaagctacgtacatggcctgttcc
 atagagtagtctataagaaattgatgacttgattcatccgaatggctggctgtaacacctgggttacgcatga
 acacctcttttcagttgtctcaagacacctttctttctgtacttatcagacaaggactgaaaggcagagac
 tgctactgttagacattttgagtcagcgttttccctggacatagctttgtcatgaaagccctttacttctga
 gaaacttctagcttcagacacatgccttcaagatagttgttgaaacaccagaagaaggagcatggcaatgc
 cgaaaacacctaagataataggtgaccttcagtggtggcttcttgcag**AATCCAGAGAGACAGACTTGCTCA**
GTGGGATGGATGGCAAAGGGTCTTACCAGGTGAAAGGCTGGAAACTTACACAGCAACAGTTTGTGGCCCTTT
TGTGGAAGAGACTGCTAATTGCCAGACGGAGTCGGAAGGATTTTTTGCTCAGgtgagacgtgctgttttcg
 ccagagactctggcttcatgggtgggctgcaggctctgtgaccagtgaaggcaggatagcatcctgggtcaag
 atatggatgccggagccagatttatctgtatttcaatcccagttctattccttggcagttgtgtatccgctg
 gcaagttacttctctatgcctcaatctcctcatctgtaaaaatggggataataattacctgcaatacaggg
 ttgttacgaaaaataaaaatgaataggtgcttagaatgggctgacattagtaagtgttagttttgtgtgt
 gtatatgttatttttatttttggaggagaacataaaaaggacaaaagtgtagaaaaactgggtgggtgtattca
 gctgtcataacatgagagttggttatgcccagatgcacttgacatgtgaatttattagaaacatgatttttct
 ctgagttgatgttttaactcaactgatagaaaagataggtcagaatatagttggccaacagagaagacttgt
 tagactattgtctgtcagtgctgtttgcatgctaacttgcttagttagaaagggttaaattttttcactcta
 taaaatcaagaaatatagagaaaaggctctgcagagagtctttcatttgatgatgtggatattgttaagagcg
 ggagtttggagcatacagagctcaagttgaatcctgactttgctacttattggctatatgacctggggcaag
 ctgcttagtctctctgatcctcagttacctttgtttgttgatgatgaccattgataacacaaccataaaataa
 tgacaacatagagatagttctcattatagtagttgttatacagaattattcactcaatgttaattttctgca
 ttgaaatcccagaacattagaattgggggcatattttgaatctttaagggttataaggaatacatttctcagc
 aataaatggaaggagttttgggttaacttataaaagtatacccaagtcattttttttcagagaagatatgggt
 agaaagtcttaggaggttgaagaaggaaattggatatttattctttctgagactatcatgggagataatgact
 atgggtgtccatgattggagccgttgctgtagagttgggtttattatagtgtaggatttgaatgggccatgt
 gttctcagacctcagattaaaawgagaaaaactgaggccagtggggagcgtgacttcacatgggtacacttgt
 gctagagacagaaccaggattcaggacttctggctcctggctcctgggttcatggccaatgtagtctttctc
 agtcttcaggaggaggaagggcaggaccagtggtctgagtcaccctgaatgtgagcactatttacttctgt
 aacttcttggcttagtgctctgccaggtggccataacctctggccttgtgttgccagagaaaaagggtttagt
 tttcaggctccattgcttcccagctgccaagaatgccttgggtgcagcacagtcataaggccctgcattcctca
 ttgccgtgctgggttgggtcggggaggtgggctggactcgtagggtatttgcccttggccttgtttctaact
 tgccgtttctgtgtccccctgccctccactgcctgggttaaag**ATTGTCTTGCCAGCTGTGTTTGTCTG**
CATTGCCCTTGTGTTTCCAGCTGATCGTGCCACCCTTTGGCAAGTACCCAGCCTGGAACCTTCAGCCCTGGAT
GTACAACGAAACAGTACACATTTGTTCAGgtatgtttgtcttctacatcccaggagggggaagattcagagcag
 accaaagatgtttacgagggccaagggaatggacttcagaattacacgggtggaatgaattttactgctgcgg
 ctcaggtccctgtataagctaatactgcatgcatagaaacagcagcgaactaacctgaataataggccagtc
 ttctgttgagcctttcagcctctctcctctcactgttgtcaggaacagccacatgtgttttaggtg
 aaataatccacccttgcaaaaatccatgattaagttataaaaatatttggatttgggagctgtgttttaatt
 ctgtaactgagtcacagggcacactgtcaaagcatagaacctccagagacttggtttctgcaaaagtataatt
 catgtaattattatctattctgttatatttgggatgttaggtagtgtttgttcttttagataaaaaatatcccc
 cactctgtaacaatacattaaatcaaagaaaaggacaaaggatttttctgggtcttgttagcaggagctttc
 ttcagtcctgaaagattttagacctgtagatgggggaactgtgtcagtgatacaaaaagggaagcattttaa
 aaaaaaaaagtatatatatatatatatatatatatatgtaatgtgaattggcctctttttctctaagzcca
 cattttzcttcttacatagttcaggttactttatttttttcttccggctgctgacctgtattgcccgtg
 gttgtggaacatagcatgtgtttgtgacctgtgcctgttatttttctgcttcttagttgtgcatgcaagag
 taaaaagttttcttgccctttcttggaatcctgtgtgtgtgccaagggaataattgtgaaagcacttt
 tgaaatacttaatgagttgattttcttcaaatataaaaaaataataatgtatatgtgtatgtacatgtgt
 gtacacatacacacctttatacacagccatttataaaacaaagctccactttggagtgctctacgtcacct
 gatccgaatacagggccagagctctgagatccttctgggtgggtttctgtgttttgttcttctgttttaag
 agcctgtcacagagaaatgcttccataaatgtttaatttataaaaaacatttttatctctcgattactgggtt
 taatgaattactaagctggctgctctcatgtaccacag**CAATGATGCTCCTGAGGACACGGGAACCTTG**
AACTCTTAAACGCCCTACCAAAGACCCTGGCTTCGGGACCCGCTGTATGGAAGGAAACCCAATCCCgtgag

cggggtggtcacacctgtaatctcagtatcttgggatgccgaggtggactgatcacctaagggtcaggaggt
cgaaaccagcttggccaatgtggcgaaaccctgtctctactaaaaattcaaaaattagccaggcatgggtggc
acatgcctgtaatcccagttacttgggaggtgaagcaggagaattgcttgaacctgggaggtggaggtgcg
agggagccaaaattgcccactgtactccagcctgagcaacacagcaaaaactccatatcaaaaaataaaatg
aataaaataacagctaattctagtcacagtaataactccagtgaacagaagatttatttaggcatagtgaatga
tggtgcttctctaaaaatctcttgactacaaagaatctcatttcaatgtttattgttttagatgttcagaataa
attcttgggaaagaccttggcttgggtgtaagtgaattaccagtgccgagggcagggtgaaccaagtctcagt
gctggttgactgagggcagtgcttgggacctgtagtcaggtttccggtcacactgtggacatgggtcactgtt
gtccttgatttgttttctgtttcaattcttgtctataaagaccctgtatgcttgggtttcatgtgatgacagA
GAAAAAAAACACTGCAGATATCCTTCAGGACCTGACAGGAAGAAACATTTCCGATTATCTGGTGAAGACGT
ATGTGCAGATCATAGCCAAAAGgtgactttttactaaacttggccctgcccattattactaattagagga
attaaagacctacaaataacagactgaaacagtgggggaaatgccagattatggcctgattctgtctattgg
aagtttaggatattatcccaaactagaaaagatgacgagagggactgtgaacattcagttgtcagcttcaag
gctgagggcagcctgggtctagaatgaaaatagaaatggattcaacgtcaaattttgccacttagtagcaactt
gaccaggtaactgggttatccttttaagccttagtttatctaaattgtgatattaatgttgctcttataagt
ttgtcatgaggactaaatataatgggtgtacatagagtgccttgggtactctctgatggggactccatgata
atttgtggtctcatggagggagctctgggaagggttaggagcctgcctggctctgcagccttgggagagcc
ttctagcttcccaggacatggcagcctagtgttgatgcttggctcagcaaatgtttgttctcgtttctctt
ccatcaacttggtcagttggggctcttccagtttaggagatctcagtgactttaaatggcatgggcatgctgg
agtgatagtgaacatgagtttctaagaaagaagcataatttctccatatgtcatcccaattgaaatattat
tgttaattgaaaaagcttctagggcagggcaggttgcctcatgacctgtaatcccagcactttaggagggccaag
gcgggtgggactcacttgaggtcaggagtttgagaccagcctggccaacatggggaaacctgtctctactaaa
aatacaaaaataagctgggctgggtggtgctgctgtaatcccagcacttgggaggtgagggcaggagaat
tgcttgaatctgggagggcggaggttgagtgagctgagttcatgccattgcattccagcctgggcaacaaga
gcgaaaccatctcccaaaagaaaaaaaagaaagaaagccttctagtttgggtacatcttgggtctataag
gtggtttgtaaattgggttaaccaaggcctgggtctcatataagtaatagggtatttatgatggagagaag
gctggaagaggcctgaacacagggcttcttctctagcacaccctacaaggccagctgattctagggttat
ttctgtccgttcttataatcctcaggtggatatttactccttttgcacatttaggaatagggtcagtgcttt
ctttgaactgatttttgtttcttctgtctctgcagCTTAAAGAACAAAGATCTGGGTGAATGAGTTTAGgtaa
gttgctgtctttctggcacgtttagctcagggggaggatgggtgttgtaggtgtzcttggattgaagaaagcc
ttggggattgtttgtcactcacacacttgggtgcatctcactgtgaggaggacagaagccctgtgaaca
tgtggagcacacagggggcacagacagatttagattaggcctgctttatagagtttctgcctagagcatcatg
gctcagtgcccagcagccctccagaggcctctgaaatatttgatatactgatttcttggaggagaatcaga
aatctcctgcaggtgtctagggatttcaagtaagtagtgtgttgagggggaatacctacttgtactttcccc
caaaccagattcccagaggttcttaaggactcaaggacaatttctaggcatttagcacgggactaaaaaggt
cttagaggaataagaagcgccaaaaccatctcttgcactgtatttcaaccatttgccttctgggtttt
gaaggaacaggtgggactggggacagaagagttcttgaagccagtttgtccatcatggaaaatgagataggt
gatgtggctacgtcagggggccgaaggctccttgttactgatttccgtcttttctctgccttttcccc
agggccaggaccctggatctctgggcagagcagagcagggccctataatagccctcatgctagaarggag
cggagcctgtgtataaggccagcgcagcctactctggacagtgaggggttccactctcccaactccccat
ctgcttgctccagaccacattcacacmcgagccactgggttggaggagcatctgtgagatgaaacaccat
tctttcctcaatgtctcagctatctaaactgtgtgtgtaatcagggccaggtcctcctgctgggcagaaacca
tgggagtttaagagatttgccaacatttatttagaggaagctgacgtgtaacttctzzgaggcaaaatttagccc
tcctttgaacaggaatttgactcagtgaaacctgtacacactcgcactgagtgctgctgctgatgatactgtg
caccctcagtgctgggttttaattgtcaggtgttcttttagGTATGGCGGCTTTTCCCTGGGTGTCAGTAAT
ACTCAAGCACTTCTCCGAGTCAAGAAAGTTAATGATGCCATCAAAACAAATGAAGAAACACCTAAAGCTGGCC
AAGgtaaaaatatctatcgtaagatgtatcagaaaaatgggcatgtagctgctgggatataggagtagttggc
agggttaaacggatcacctggcagctcattgttctgaatatgttggcatacagagccgtctttggcatttagc
gatttgagccagacaaaactgaattacttagttgtacgtttaaaagtgtaggtcaaaaacaaatccagaggc
caggagctgtggctcatgcctgtaatcctagcactttgggaggtgaagcgggtggatcacttgaggtcagg
agttcgagaccagcctggcctacatgacaaaacccgtatctactaaaaatacaaaaaaattagctgggctt
gggtggcacacacctgtaatccagctacttgggaggtgagggcaggagaattgcttgaacctgttaggaaga
gggtgtagttagccaagatcgaccgttgactccagcctgggcaacaagagcaaaacctcatctcaaaaaa
caaatataatccagagatttaaaagctctcagaggttggcgcggttgcacctgttatccagcattt
tgggatccgagggcgggcaaaagcacaaggtcaggagtttgagaccagcctggccaacatagtgaacacctgt
ctctgtcaaaaaacatagaaaaattagccggcatgggtggcgtgcgctgtaatcccagctactcgggaggt
gaggtgagagaatttcttgaacccgggagggcggaggttgagtgagcccagattgcaccactgcactccagc

ctgggacagagcaagactccatctcaaaaaagctctcagaacaaccaggtttacaaatttggtcagttg
gtaataaactgggtttcaacatactttgctgaaayaatcactgactaaataggaaatgaatctttttttt
ttttttttaagctgggaagctggctgtgtaggacctgataagtaactcatttctctgtgtctcaggtt
tccatttttaggtgagaattaaggggctctgataaaacagaccctaggattgtggacagcagtgtrtagtcc
tagagtccacaagctctgcttttgagtgtgggcccagtgatctggcacatctgcaggcagagcgtgggtctg
gctcttcagatgatgccggtggagcactttgaggagtcctcaccacccgtgataaccagacattaaaatct
tggggctttgcatcccaggatttctctgtgattccttctagacttgtggcatcatggcagcatcactgctgt
agatttctagtacattgggttctcaggagccgtttatttaattggcttcacatttaatttcagtgaacaaggta
gtggcattgctcttcacagggccgtcctgtgtccacaggttccagattgactgttgccccttatctatgtg
aacagtcacaactgaggcaggtttctgtgtttacag**GACAGTTCTGCAGATCGATTTCTCAACAGCTTGGG**
AAGATTTATGACAGGACTGGACACCARAAATAATGTCAAGgtaaaccgctgtctttgttctagtagctttt
gatgaacaataatccttatgtttcctggagtaactttcaactcatggttaaagttggcaggggcattcacaaca
gaaaagagcaaaactattaactttaccagtgaggcagtaacggtgtagtgtagtgtattcagagaatttgcttg
ccaccagacataccaggtaaccttgactaagttacttaacctatctaaacctcagtttccatctgtgaaa
tggagacagtaatcatagctatttccaaactgtgtgagaattcaatgagttaaagggtataagggtcctcacc
acagcgccctgcccacatagtcagtgatcactatgtcctgaacactgttaattacttcgccatattctctgatc
atagtgttttgccctggatgtgactagaatttcttctgaggtttatgggcatggttggtgggtatgcacc
tgccctgcaggagcccggtttgggggcattacctgttacctggatgttttcttctcag**GTGTGGTTCAATAAC**
AAGGGCTGGCATGCAATCAGCTCTTTCTGTAATGTCAACAATGCCATTCTCCGGGCCAACCTGCAAAAG
GGAGAGAACCCTAGCCATTATGGAATTACTGCTTTCAATCATCCCCCTGAATCTCACCAAGCAGCAGCTCTCA
GAGGTGGCTCTgtaagtgtggctgtgtctgtatagatggagtggggcaaggagaggggttatggagaagggg
agaaaaatgtgaatctcattgttaggggaacagctgcagagaccgttatattatgataaatctggattgatcc
aggctctgggcagaagtataagtttacgaattggctgggtgggtcttctgaactgcagaagagaaaatgac
actgatatgtaaaaaatcgtaacatttagtgaattcatataaagttagttcaaaaattgttaattaaattata
atttaattataagtgtttaatcagtttgattgttttaaaaaccactgttttaaaatttggtggaatatgtttt
tattagcttgtatctttaattcctaaattaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
gtgtgtgzztgtaagtttaaaagccaggatgagctagtttaagtagtcagcccttgagtcatacagatct
gggtttgaatctggtctctaaactttatagatgtatgatattaaatgaggcagttcatgtaaattgccaagc
ccagcactcagcacagagttgatatttcacacacattagatacctttcctgtatgtggagcatggcagttcc
tgtttctgctttactcctacaggataactaatataggacactaggatctttataccaagaccccatgtaatgg
gcttatgagaccattcttcttataaaaaatctgacagaatttttgtatgtgttagatcaataggctgcatact
gttattttcaagttgattttacagccagaaatattaattttatttgagttagttacagagtaaatatttctgctc
catttagttttcaagccccactagtcctttgtgtgtgaaaaatttacaacttactgctcttacaaggctcatga
acagtggaccaaagtgaatgccattaaccactctgacttccttcattagttttattgtgacagtggaactctt
ttgacctcagtaataaccagtttggcattttacattgtcatatttttagacttaaaaaatgatcatcttaaccct
gaataaaatgtgtctggtgaacagatgtttttccttgggctgtgcctcagatatctctgtgtgtgtgtacgt
gtgtgtttgtctgtgtgcatgtcctcactgattgagccctaaactgcatacaagacccctcagattttcac
acgctttttctctccag**GATGACCACATCAGTGGATGCTCTTGTGTCCATCTGTGTATCTTTGCAATGTCC**
TTCTGCTCCAGCCAGCTTTGTCTGATTTCTGATCCAGGAGCGGGTCAGCAAAGCAAAACACCTGCAGTTTCATC
ATGTGGATGTAAGCCTGTCTACTTGGCTCTCTAATTTTGTCTGGGATATGgtaaggacacaggcctgtgt
atctttctgatgtctgtcagggccatggattgatatggataagaaagaaagagctctggctatcatcaggaa
atgttccagctactctaaagatgtatgaaaaagaaatagccagaggcaggtgatcactttcatgacaccaa
cacagcattgggtaccagagttcatgtcacaccagaggggaaaattctgtacacaatgatgaaaattaatacc
actaccacttaagttcctatgtgacaactttccaagaatcagagagatacaagtcaaaactccaagtcaat
gcctctaacttctctgatgggttttaacctccagagtcagaatgttctttgcccacttaggaaagccatctg
tcatttztgaaaactctgtacattttatcagcagcttatccatccattgcaaataztgtttttgtgcccagcca
caatataattgcttctatttggaccaatatgggggatttgaagggaattctgaagttctaatatatttcaact
ctactttacaatatctccctgaaatataatctccctgtaacttctattaattataagctacacagagcaaatc
taattcttctcccaccgaacaagtccttgatatttaaaaaataactctcatactctcatttaacctgagtat
taccagataagatgatataatgagaatacaccttgtaacctccgaagcactgtacaaatgtgagcaatgatg
gtggagatgatgatgatctttgtgtttataccaagccccttagactgtgtcactctctgtatccggttg
tccttgtatggccatgctgtatattgtgaatgtccggttttcaaaagcaaaagccaagaattaacctgtgtt
caggctgtggtctgaatgggttatgggtccagaggagtgatcttttagctcacacttctattactgcagcac
aaagattttgcatttttgaaggagcaccgtcttactggcaacttagtggttaaaccaaaacctccatttcaca
caaatgattgtgaaattcgggtctccttcattctatacaaaattcatttgatttttttgaactaaactttat
atztatccatattaaattacatgggttttattttgttttatcttgattcagtaattactcctttcagtaaa
cacagactgagtgctgtgtgtgtgacttatgccaggcataggtgattcagagatgaaagggtcaagtcctga

accatctcttgtcttcctgggtattatctgtccctccctgcttttagagctcctgaaatttgc tagaagcat
gtcttcatctaagttgttgataaacacatcaagtaggattggactgaggcagagccctgtagtctgaagctg
cagttcttctagcggctgacaagcccactatcacttccctgctggtgctttgctctgccagctgtgaattc
tcataattgtcctatcgtaagctctttatttctgcattttactgcttgatacactgtcaggacagactttaa
aattattctcagtgccgatgaaacaattctgacattcatgttatgagcagttacctcataaatagattacatg
tgagattgaacttgggcagactataatatagcattaatgatgaaacagacacagtcattctcggaagaaga
atagaggcttatttgcctgtgaaattaaaattactctgactgggaatccatcgttcagtaagtttactg
agtgtgacaccttggcctgactgttgaaagacagaaagggcatgtagttataaaatcagccaaggggaaa
atgcttgtcaaaatgtattgtcgggtattttgattaatagtttatgtggcttcattaattcagagttactct
ccaatatgtttatctgccccttcttctgtctgataatgggtgaaaacttgtgtgatgcattgtatatattgattta
gggggtgaactggatgtcttctgttttctcatttttag**TGCAATTACGTTGTCCCTGCCACACTGGTCATTATCAT**
CTTCATCTGCTTCCAGCAGAAGTCTATGTGTCTCCACCAATCTGCCTGTGCTAGCCCTTCTACTTTTGCT
GTATGGgtaagtcacctctgagtgagggagctgcacagtggaataaggcatttgggtgccagtgtcagaagga
gggcagggactctcagtagacacttatcttttgtgtctcaacag**GTGGTCAATCACACCTCTCATGTACCC**
AGCCTCCTTTGTGTTCAAGATCCCCAGCACAGCCTATGTGGTGCCTCACCAGCGTGAACTCTTCATTGGCAT
TAATGGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTACCCGACAATgtgagtcagtcagagagaacactcct
gctgggatgagcatctctgggagccagaggacagtggttaattgtgatcttattccacttgtcagtggttatt
gacactgctgactgccttgtcctgtcttcagagtcgtcttccctgagaaggcaaacacttcttcttctg
ctgtgccttacatttctgtgtgcaagccttccagtttcttcttgcagagtttttttacttcttcttcttcttca
atgttgcctttaccacagagtagctcctctgccttccactttacacatgagagctgggcgacgzcattcagtc
ctaaggctctttaccatcacctctcttgggtgttttattgtcatctctaagatcaatgcctttagccttgatc
ataaccttgaactctaatctcaaattctcacttgcctagtggattgctccatttagatagtatatagatacc
ccaacctggatatgtcctagttttcttctcccttggaaacttaatgcttttcttgccatccctgtcacactca
gtggcactaccatccactcggttggcccaagctggctcttagagttatcctagatgcttgccttctgtgtgca
gatttccacattcaactggttatgttgcagttcttccaggtatggacctctaaaataaggcttctctctcc
attccggttgtcattgccttctgctcaaacacagcacacaaggccttttacagttgcacaactcttctctgctc
ataccaccacaccccttctccagctgtaagcttcagatgagttgcctccaaccaccatgctcctgtaggcct
ggcttgaaatgcccttcttctgtcacaggggtctggtagtatatcccttgccttcaagatttagctaaaatg
tgaagcttcttaccctgctgggaggtgttctctcttctctctgtgctctcagagtccttagtccatgcctc
cagtacaacgtacatccacttacatggtaatttctctgtttacatacttttctactcggagtgaggctgtt
tcttaataatttctgctctcccatgccctagcacagtgcatccagcgtatagccccttattcagttggtaga
tatttggccactgttgccttgtgggatcataagttctgatgtatttgagaagaatttctaaaattctgacaa
aatcctgaaactcaaatattgacccagacatgagcaatttgcctttcaaatgctaagggtttttaaaggat
ttgctttaattaaatctagcctgtttctaaagctttattcattatttctccatactcagagcattttctcaga
ttttctaaagaatagaattttattgtcatatatcatcagctatgcctgctgctattttaattggtatctgaat
taaaaggctctggttgtccctagagaatcaaattttttcttccactcccatatttcagaacttgatacatttt
taggataaacatgaatgacacccggttcttctccctcaccctcccttccctccattttttttttttttttt
tttttag**AAGCTGAATAATCAATGATATCCTGAAGTCCGTGTTCTTGATCTTCCCACATTTTGCCTGGG**
ACGAGGGCTCATCGACATGGTGAAAAACCAGGCAATGGCTGATGCCCTGGAAAGGTTTGgtgagtgaaagcag
tggtctgtaggatgctttaatggagatggcactctgcataggccttgggtaccctgaactttgttttggaaaga
agcaggtgactaagcacaggatgttccccacccccatgccagtgacagggctcatgccaacacagctggt
tgtggcatgggttttgtgacacaaccatttgcctgtgtctctgatagcattgagaaaagtgaaggggcagtt
ttgaaggtaaggaaaatagtggtatttgcctggatccactggctcatgccactgtctgggttgggttagaagc
actggaaaagtcaaaccataactttgagaattaggtgatcagggaaatcagaaggaaagatgcaaaccttggc
tcttttaggcgaatcatgtgcctgcagatgaggtcatttattatcttttacacagtcataaaaattataatg
tattacatcttttctacctttagaatggttaaaaaatatttctccggtagccatatgattattattcatcca
ttagataatatagtcaaatgggcatgttatttactgttcatagaagaggggcttttttgcaacttgggctac
aaaggagatatgtaaggaaatttaaggaaatggttacatggaactagatttaattgaatctagtggtttaattg
attcactaggatatatgctactgaaaggggaatctgcttaaagtgccttctgatatttattattactaaaac
ttagaatttattaaaaaactgactgtgaaaattacttgggtcggttgcctttttaaaggatttttggcat
gtctcattaaaaaaaagaataactagatatcttcagtgaaagttaaaaatcgaataacacattgggtctgaaatt
ctgattgatactgggtcataaaaagttttcccaaatcagacttggaaagtgatcactctctgttactcttt
tttcttgcctgaggtgatagccatttgccttatttggaaagatcggtgaattttaaggaaacataggccaaa
tttgaggaaaggccatggtttttagatccctccattctgaccggatctctgcattgtgtctactag**GGGAGAA**
TCGCTTTGTGTACCATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCCATGGCCGTGGAAGGGGTGGT
GTTCTTCTCATTACTGTTCTGATCCAGTACAGATTCTTCATCAGGCCAGgtgagctttttcttagaacc
gtggagcacctgggttaggggtcacagaggaggcgacaggggaaacactaccaatgggggttgcatgaact

gaactcaaaatgatgtataaaactgattttcctgatgtgggcatcccgagccccctccctgcccacctgg
agactgtggcaagtaggttttataataactacgttagagactgaatctttgtcctgaaaaatagtttgaaagg
ttcatttttctgtgtttttccccaag**ACCTGTAAATGCAAAGCTATCTCCTCTGAATGATGAAGATGAAGA**
TGTGAGGCGGGAAAGACAGAGAATTCTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGGAGTTGAC
GAAGgtgagagagtacaggttacaatagctcatcttcagttttttcagctttatgtgctgtaaccagcag
tttgctgacttgcttaataaaagggcatgtgttcccaaatgtacatctataccaaggttctgtcaatttta
ttttaaaacaccatggagacttcttaagaattcttactgagaattctttgtgatatgaattccattct
cgaatacattgggttttatatgcttacatttatgtgttagttattaaaacatactaataattgtatatctagtc
aaaactgaggtagagagaataaatgggtgattttgagtttgagtttcatagtcocaaaagctgatataattgc
ctgtgttcaagagggcttatatcagccctctagatgccagcatctccaaattttacttttttggaaatctgta
cagtatttgcaatatttttattacaaatttctactctgtggaatttaatttttaaaatacctgcaatacata
tatatgttgaatagatgaaaaattatgtagatratatgaatgatacgggttctaaaaagacaggttaaaaag
taagttcactttttattttgagcttcagaatcattcagaagccagtcgccacaaacgcagaccaaggtcttg
gcacatcaaatatgcctatggccttagggttattgacaagtcttatgttgagtgatgtggtttatagtcct
gccttccacagttgcttgggagagctgtgagtcactgaggttatgaatgtttacattttgtttgttgca**A**
TATATAGAAGGAAGCGGAAGCCTGCTGTTGACAGGATTGCGTGGGCATTCTCTCTGGTGAGgtaagacac
tttgtctatattgctgtttgtccctattagttcagactatctctacccaatcaagcaacgatgctcgttaaga
ggtaaaagtggatttttaaggcttctgtatttatgccaggttgagcaattagtcacgcgagaagagagggac
cctgtatgtcaagagaatgatttcagagaatccaatacaatttaagaaaaagcatggggctgggagcagtg
ttcactcctgtaatcccagcactttgggagggcaggtgggaggactcacgaggtcaggagattgagaccat
cctggccaacatggtgaaaccccatctctactataaatacaaaaaattagctgggcatagtagtgcatctctg
tagtcccagctactcgggaggtgagggcaggagaattgcttgaacctaggagggggaggttgcccagattgc
gctgtgcactccagcctggtgacagagtgcagactcatgtcaacaacaaaaacagaaaaagcacgcacatct
aaaacatgctttttgtgatccatttgggatgggtgatgacattcaaatagtttttttaaaatagattttctcct
ttctgggtttccgtttgtgtttcttttatgcccttttgccagagtgggtgtczaatttggtazgctggcct
tcattactgtttttcaczaczaattaacztttggcctcaacttgacaactcaaataatattataatacacgc
cacacttaaaatgggtccattatgaaatacatatttaaatatctatacagatgtgttaaaaccaagaaaaat
ttgattcttctctgatatttaagaattgaaggtttgaggttagttacgtgttaggggcaatttatattcatgtt
tttagagtttgcttatacaacttaattctttccttttcag**TGCTTTGGGCTCCTGGGAGTTAATGGGGCTGGA**
AAATCATCAACTTTCAAGATGTTAACAGGAGATACCACTGTTACCAGAGGAGATGCTTTCCTTAACAAAAAT
AGgtgagaaaaagaagtggcttgattttgctgcaaagactttgttttttaatttttaagaaaaataggttg
tatttttgattacagtggtatttttagagttcataaaaaatgttgaaatatagtaaaagggtaaagaagcacat
aaaatcatccatgatttcaatatctagagataatcacaatttacatttcccttcagctcattctctctctt
taacagctttattcaggtataatttacatacaatataatttgcttggtttttaagagtataatttagtgatt
tttggtaaattgagagttttgcaaccatcaccacaatccagttttagaacttttccatcaccacacatctgt
cttatatacacatataaatgtgccatacaattgagatcatactgtatgtagaatttaaaattagtttttatt
gttaatgagtgattatgaatatctccagtggtttacatttccctaagatgtggaattttacattgctacat
aaaatccccctatgtacatgtacataatttataaataaaattccttataaatgttggacacattagtttc
catttttctactatgtaaatatgtccctgtatacatcttttattatttctcaggaacaattcctacaaagta
aattggcctctctaaagagcatacaaattgactgagccacggttagggcattttctgagactgcacaggtca
caaagcaatctgatctttgggaatacagctacattttataggcttcttagataatgttactctaagtaatt
aaatatgtggggtctctctgggcttttttttttttgagacggagtttctacttactgcccagggtggagag
caatggcgcgaccttggtcactgcaacctccgctccaggttcaagcgattctcctgcctcagcctcctg
agtagctgagattacaggtgcccgcacaatgcctgcctaattttttgtattttcagtagagatggggtt
caccatgttggccagactggtctcgagctcctgacctcaggtgatccacctgcctcagcctccaaagttct
gggattacaggtcatgagccactgcgcccggcttctctggacttattatgtggagagatagtacaaggcagtg
gctttcagagttttttgacctgacctgtgtgggaaatacattttatatctcaacctagtagtacacacag
acatgtagacacatgtataacctaaagtttcataaaagcagtaacctactgttactaattgtagtgcactctgc
tatttcttattctaccttatactgcgtcattaaaaagtgctggtcatgaccactaaattttatttcccaaa
ccactaatgaacaatgactcacaatttgaacacactggacagggggatagccaataaaattgaaaagagcaa
ggaaattaatgtattcatgatctctctcctgtctcttacatttttgagtagcaatgtaaaaggaatcctaa
gagaacagacattctgggaatagcaggcctagcgctgcacaactgcttcttaggttgcctcctagtaccaa
gctcctgacgcataagcagtgagcagtaataaccagcccatagtaaggtttgtcacagggactgggtgtaag
aactgattttgggttgatagctgtgagggcctggcagcgtgtccacgtgtgctcaatcctaattctgaaaa
aggctgacctgggggtgctaattagatacacagagaggaatgaatgctgccagaaggccaagttcatggca
atgcccgtgtggctgaggtgcagtcacagtcgtggaacgtgaacactgaacttctctcacatgtgattcttc
acttgactggcttcatagaaccccaagccaccaccacacataaattgtgtctctaggttctgtgttg

tcacactcaaaatttctgggccttctcatttgggtgcatgtgaatgggtgcatatgagtggaagtctaggatggg
gccttagcggttaaagccctgggtagtgtagtgagattgttggttaaagaatgtgcagtggttggcatgacc
tcagaaattctgaaatgggactgcacctgcagactgaagtgttcagagagccagggaggtgcaaggactggg
gagggtagaggcaggaaccctgcctgccaggaagagctagcatcctgggggcagaaaggctgtgctttcaag
tagcagcagatgtattggtatccttgaatggagaagcatactttacaggaacattaggccagattgtctaa
ccagagtatctctacctgcttaaaatctaagtagttttctgtcctttgcag**TATCTTATCAAACATCCATG**
AAGTACATCAGAACATGGGCTACTGCCCTCAGTTTGATGCCATCACAGAGCTGTTGACTGGGAGAGAACACG
TGGAGTTCTTTGCCCTTTTGAGAGGAGTCCCAGAGAAAGAAGTTGGCAAGgtactgtggggcacctgaaagcc
agcctgtctcctttggcatcctgacaatatataccttatggcctttccacacgcattgacttcaggctgttt
ttcctcatgaatgcagcagcacaataatgctggttctttgtatctgctttcaggggtggaaacctgtaacggtg
gtggggcagggctgggtgggcagagagggagtgctgctcccaccacacgagtcctcttccctgctttggct
cctcaccagttgtcaggttatgattatagaatctagtccactcagtgaagaactttcatacatgtatgtg
taggacagcatgataaaattcccaagccagaccaaagtcaaggtgctttttatcactgtag**GTTGGTGAGTG**
GGCGATTTCGGAAACTGGGCCTCGTGAAGTATGGAGAAAAATATGCTGGTAACTATAGTGGAGGCAACAAACG
CAAGCTCTCTACAGCCATGGCTTTGATCGCGGGCCCTCTGTGGTGTTCGTGgtgagtataactgtggatgg
aaaactgttgttctggcctgagtggaataacatgactgttcaaaagtcctatatgtccagggctgttgatga
ttggcttgtcttccccagggacagcagagcaaccttggaagcagaggggaagcttctcccttggcacaca
ctgggggtggctgtaccatgcctgcagatgctccaaatagaggcactccaagcactttgtttcttagcgtga
ttgaggctggatgtgatttgatcttctctggaacattcttctaatacatctttgtgttcattccctgaa
aatgaagagtgtggacacagcttttaaaatccccaaaggtagcaactaggtcatagttcctztacacacggata
gatgaaaaacagatcagactgggaagtgcccttgacctttttcttctgtagataagagcattgatgttat
tacgggaagaagcctttgaggttttatgtattccacctcggctctggaatttgtttctgtaaggctaacagt
tgcaatatactagggtaatctgagtgagctggaattaaaaaaaaaaggaatttcaccccaatcttatactg
acttcaatagaggtttcagacaaaaagtgtttttgtatatacttatcagtcagtaaaagataattacaacta
aatggcctttttccctccctattttatttggagaaatttaattacataaaaaagttactcagaatatttgagtt
tctgcacataaagacattttataataatgacctgttttacaataatgaatttgaaagttactctaattctttg
attcatcaagaaataactagaatggcaagttaaaatttaagctgtttcaaagatgcttctgcatttaaaaaac
aaatttatctttgattttttttccccccagcaataagacttattttattctaattacag**GATGAACCCACC**
ACAGGCATGGATCCCAAAGCCCGCGGTTCTTGTGGAATTGTGCCCTAAGTGTTCAGGAGGGGAGATCA
GTAGTGCTTACATCTCATAGgtccgtagtaaaagtcttgggttccctcactgtgggatgttttaactttccaag
tagaatatgcatcattttgtaaaaatttagaaaatacagaaaagcaagagtaaaacaattattacctgaaa
ttatatatgcatattcttacaaaaatgcaagccagatataaatactgctctttttcacttaatatattgtaa
acattattccaagtgcagtgcattttaggtgtcatttcttatagctggatagattccattaggtatatactctt
atttaactattccccctttttagacatttggattatttccaacttgttcacaattgtaaacaccactacac
tgaaacagcatcatccctatatccacatgtacttgttaacagaatacaattccctaggaagctggaatgctgga
agtcatgggtgatgttctcatggttacagagaatctctctaaaaactaaaacctctttctgttttaccgcag**TA**
TGGAAGAATGTGAAGCTCTTTGCACCTAGGATGGCAATCATGGTCAATGGAAGGTTTCAGGTGCCCTTGGCAGTG
TCCAGCATCTAAAAAATAGgtaataaagataatttctttgggatagtgccctagtgagaaggcttgatattta
ttcttttgtgagtataataatggtgcctctaaaaataaagggaataaaaactgagcaaaacagtatagtggaa
agaatgagggctttgaagtccgaactgcattcaaattctgtctttaccatttactgggtctgtgactcttgg
gcaagttacttaactactgtaagagttagtttccctggaagatctacctcctagctttgtgctatagatgaa
atgaaaaaaatttacatgtgccagtagtggtgagagcgcaagctttggagtcacacacaaatgggtttgcat
cctggccctaccaattatgagctctgagccatgggcaagtgactaactccctgggcctcagtttctctgtaa
catctgtcagacttcaggtccaggtgaggattaaaggagatcatgtatttacagcacatggcatgggtgct
tcacataaaataagtatatttagtaaatgataactgggtccctctctcagaaacttatttctgggcctgccagg
ggccgcctcttttcatggcacaagtgggttcccagggttcagttattcttttaaatagttttctgggatcc
tccatttgggtatttttccctgcttccag**GTTTGGAGATGGTTATAACAATAGTTGTACGAATAGCAGGGTCC**
AACCCGGACCTGAAGCCTGTCCAGGATTTCTTTGACTTGCAATTTCTGGAAGTGTTCATAAAGAGAAAACAC
CGGAACATGCTACAATACCAGCTTCCATCTTCATTATCTTCTGCCCAGGATATTACAGCATCCTCTCCAG
AGCAAAAAGCGACTCCACATAGAAGACTACTCTGTTTCTCAGACAACACTTGACCAagtaagctttgagtg
caaaacagatttacttctcaggggtgtgattcctgccccgacactcccgcctataggccaagagcagtttg
tatcttgaattgggtgcttgaattcctgatctactattcctagctatgcttttactaaacctctctgaacct
gaaaaggagatgatgcctatgtactctataggattattgtgagaatttactgtaataataaccataaaaaac
taccatttagtgagcacctaccatgggccaggcattttacttgggtgcctaatacctattttaaattagataaaa
aagtaccaaataaggtcctgacacttaagaagtactcagtaaaatatttcttccctcttccctttaaataaga
ccgtatgtgcaaaagtaaatggatgactgagcagttgggtgatgtaggggtggggggcgatatagaaagtcag
tttttggccgggctgggtggctcatgcctgtaatcccagcactttgggagggctgaggagcagggcagatcatg

aggtcaggagatccagataatcctggccaacagggtagaaaccccgctctctactaaaaatacaaaaattagct
 gggcatgggtggcgacactttagtcccagctacttgcgaggtgagggcaggagaattgctcgaacccagga
 ggtggaggttacagttagccaaggtctcgccactgcactccagcctggggacagagcaagacccatttcaa
 ggggggaaaaaaagtctatttttaagttgttattgctttttcaagtattcttccctccttcacacacagtt
 ttctagttaatccatttatgtaattctgtatgctcctacttgacctaatttcaacatctggaaaaatagaac
 tagaataaagaatgagcaagttgagtgggtatttataaagggtccatcttaattcttttaacagGTATTTGTGAA
 CTTTGCCAAGGACCAAAGTGATGATGACCACCTTAAAGACCTCTCATTACACAAAAACCAGACAGTAGTGGA
 CGTTGCAGTTCTCACATCTTTTCTACAGGATGAGAAAGTGAAAGAAAGCTATGTATGAAGAATCCTGTTTCAT
 ACGGGGTGGCTGAAAGTAAAGAGGAACTAGACTTTCCTTTGCACCATGTGAAGTGTTGTGGAGAAAAAGAGCC
 AGAAGTTGATGTGGGAAGAAGTAACTGGATACTGTACTGATACTATTCAATGCAATGCAATTCAATGCAAT
 GAAAACAAAATTCATTACAGGGGCAGTGCCCTTTGTAGCCTATGTCTTGATGGCTCTCAAGTGAAAGACTT
 GAATTTAGTTTTTTTACCTATACCTATGTGAACTCTATTATGGAACCAATGGACATATGGGTTTGAATCA
 CACTTTTTTTTTTTTTTTTTTTGTTTCTGTGATTCTCATTGGGGTTGCAACAATAATTCATCAAGTAATCATGG
 CCAGCGATTATTGATCAAAATCAAAAGGTAATGCACATCCTCATTCAAGCCATGCCATGCCAGGAGAC
 TGGTTTCCCGGTGACACATCCATTGCTGGCAATGAGTGTGCCAGAGTTATTAGTGCCAAGTTTTTCAGAAAG
 TTTGAAGCACCATGGTGTGTCTGCTCACTTTTGTGAAAGCTGCTCTGCTCAGAGTCTATCAACATTGAATA
 TCAGTTGACAGAATGGTGCCATGCGTGGCTAACATCCTGCTTTGATTCCCTCTGATAAGCTGTTCTGGTGGC
 AGTAACATGCAACAAAAATGTGGGTGTCTCCAGGCACGGGAAACTTGGTTCCATTGTTATATTGTCTCTATGC
 TTCGAGCCATGGGTCTACAGGGTCTCCTTATGAGACTCTTAAATATACTTTAGATCCTGGTAAGAGGCAAAAG
 AATCAACAGCCAACTGCTGGGGCTGCAAGCTGCTGAAGCCAGGGCATGGGATTAAAGAGATTGTGCGTTCA
 AACCTAGGGAAGCCTGTGCCCATTGTCTGCTGACTGTCTGCTAACATGGTACACTGCATCTCAAGATGTTTAT
 CTGACACAAGTGTATTATTTCTGGCTTTTTTGAATTAATCTAGAAAATGAAAAGATGGAGTTGTATTTTGACA
 AAAATGTTTGTACTTTTTTAATGTTATTTGGAATTTTAAAGTTCTATCAGTGACTTCTGAATCCTTAGAATGGC
 CTCTTTGTAGAACCCTGTGGTATAGAGGAGTATGGCCACTGCCCCACTATTTTTATTTTCTTATGTAAGTTT
 GCATATCAGTCATGACTAGTGCCTAGAAAGCAATGTGATGGTCAGGATCTCATGACAATTATATTTGAGTTT
 CTTTCAGATCATTAGGATACTCTTAATCTCACTTCATCAATCAAATATTTTTTTGAGTGTATGCTGTAGCTG
 AAAGAGTATGTACGTACGTATAAGACTAGAGAGATATTAAGTCTCAGTACACTTCTGTGCCATGTTATTCA
 GCTCACTGGTTTACAAATATAGGTTGTCTTGTGGTTGTAGGAGCCCACTGTAACAATACTGGGCAGCCTTTT
 TTTTTTTTTTTTAAATTGCAACAATGCAAAAGCCAAGAAAGTATAAGGGTCACAAGTCTAAACAATGAATTC
 zTTCAACAGGGAAAAACAGCTAGCTTGAAAACTTGCTGAAAAACACAACCTTGTGTTTATGGCATTTAGTACCT
 TCAATAAATTGGCTTTGCGATATTGGATAACCCATTAAATCTGACAGTCTCAAATTTTTTCATCTCTTCAAT
 CACTAGTCAAGAAAAATATAAAAAACAATACTTCCATATGGAGCATTTTTCAGAGTTTTCTAACCCAGT
 CTTATTTTTCTAGTCAGTAAACATTTGTAAAAATACTGTTTCACTAATACTTACTGTTAATCTGCTTGGAG
 AAAAGAAAAATATGAGAGAACTATTGTTTGGGGAAGTTCAAGTGATCTTTCAATATCATTACTTAACCTTTC
 CACTTTTTCCAGAATTTGAATATTAACGCTAAAGGTGAAGACTTTCAGATTTCAAATTAATCTTTCTATATT
 TTTTAAATTTACAGAATATTATATAACCCCATCTGATAAGGAAAAAATGATTGTTTTAGAAGTTAAAGTC
 AATATTGATTTTAAATATAAGTAATGAAGGCATATTTCCAATAACTAGTGATATGGCATCGTTGCATTTTAC
 AGTATCTTCAAAAATACAGAATTTATAGAATAATTTCTCCTCATTTAATATTTTTTCAAATCAAAGTTATGG
 TTTCTCATTTTACTAAAAATCGTATTCTAATTTCTTATTATAGTAAATCTATGAGCAACTCCTTACTTCGGT
 TCCTCTGATTTCAAGGCCATATTTTAAAAAATCAAAAGGCACTGTGAACTATTTTGAAGAAAACACAACATT
 TTAATACAGATTGAAAGGACCTCTTCTGAAGCTAGAAACAATCTATAGTTATACATCTTCATTAATACTGTG
 TTACCTTTTAAAATAGTAATTTTACATTTTCTGTGTAAACCTAATTGTGGTAGAAATTTTTTACCAACTC
 TATACTCAATCAAGCAAAATTTCTGTATATTCCCTGTGGAATGTACCTATGTGAGTTTCAGAAATTTCTCAA
 ATACGTGTTCAAAAATTTCTGCTTTTGCATCTTTGGGACACCTCAGAAAACTTATTAACAACCTGTGAATATG
 AGAAATACAGAAGAAAAATAAAGCCCTCTATACATAAAATGCCAGCACAAATTCATTGTTAAAAAACAACCA
 AACCTCACACTACTGTATTTTCATTATCTGTACTGAAAGCAAATGCTTTGTGACTATTAAATGTTGCACATCA
 TTCATTCACTGTATAGTAATCATTGACTAAAGCCATTTGTCTGTGTTTTCTTCTTGTGGTTGTATATATCAG
 GTAAATATTTTCAAAGAGCCATGTGTCTGTAATACTGAACCACTTTGATATTGAGACATTAATTTGTAC
 CCTTGTTATTATCTACTAGTAATAATGTAATACTGTAGAAATATTGCTCTAATTTCTTTCAAATTTGTTGCA
 TCCCCCTTAGAATGTTTCTATTTCCATAAGGATTAGGTATGCTATTATCCCTTCTTATACCCTAAGATGAA
 GCTGTTTGTGTCTCTTTGTTTCATCTTGGCCCTCATTCGAAGCACTTTACGCTGTCTGTAATGSGATCTAT
 TTTTGCACTGGAATATCTGAGAAATTGMAMAACCTAGACAAAAGTTTCAACAACAGATTTCTAAGTTAAATCAT
 TTTTCATWAAARRRAAARAAAAAATTTTGTATGTCAATAACTTTATATGAAGTATTAAATGCATATT
 TCTATGTTGTAATATAATGAGTCACAAAATAAAGCTGTGACAGTTCTGTTGGTCTACAGAAattacttttg
 tgcatttgtggcaccacactactgttgaaggggtataaagccattagaaaagtagaggggaagtgatttggat
 caaaaggaaaaacttttagaaaagattcaaatgttcccttaatacataaaagagaactgaggggactacttgaa
 aataaaaggttgttttgtattttcatgttgggttaagatactgagztaacwggatttaagtgttagaggttt

tagataaatattctgcttaatgattatgaagctgactgagatttctgaaaatgctctgtagctgagcttat
ttaataaatgttcacttggtataggggaagctacaaaggcagccttcagtgctcttttgtttattcaaccaa
aaatataaggacacaatgtagcagttatactgggaaggtgctgggggtggtggcaatggtgagcaggaaggc
gaagtagatatggaacagaaatgataactaatatcggtgattccttctcttttctgtrataagtgtgtg
cagacaacatatgagcagtgctgataaatgtaaatgtattkttcatagctcattaagaatcagtttcagaaa
gagatgtctgcttatttkgctrcttgaagaatccctgtcaaacagtccttttsaggaagtacaagaggctgt
ctctattttgtgacctcaggaatggctgtgacagtgctgtgagcagtccttttctgtggcacagatctgaac
tttgtgtgcagaaaaatcttggcttcaagttagcgaagatgccccctgagcatcagcatcacaacttcatcc
tcctatcttgaagttcatgttatagtgactttaatgaaatcatagaacactgtttcttctgtgzaacaatgac
gagggagaggaaaaaactttattgaaaaataaaaaggcaggtaattagatgaaaatatgttacctatgagg
ttttgtttttgctttttgtttttgtttttgagaaacagaatctcgcctctgtcgtccaggctggagtgacg
gcatgatcttggctcactgcaacctccgcctcccggttcaagcgattctcctcagcttcccaagtagctgg
tactacaggcatgcgccaccacaaccagctaatttttgtattttttagtagagatggggtttcactatacgtt
ggccaggctggtctcaaactcctgacctaaagtgatccttctgccttgggctcccaaagtgctgggattaca
ggcatgagccaccttgcttggccctacctatgagccttgactaaaacattcttctatctgtagaaaagccca
aaagaacttttccagattcaaaaaacttggcactttgtaatggtaatgtttacattaagtaaaaaaa
aaaacccacttagcttcagtttcaagtgtttactgtgtgtgcacttcatttaattctcaacacctgc
cctatgaggtaaaaagtaccattttacatatagtaaaattacagctcagtgagataagaaactcgtccaaagg
tacaggttcagtcagtgaggaggggtctcttttggtaagttaggtatcagttaaaattgacctgtaaaa
tcacatcagcatcaatatacatttaatttaacaaatattttattgaactttactgtatgccagatacttctcta
ggtagtagggggtacaatgtagaagaaaatagaattcctgctctaataatgtttatattttagtggtgaaaga
tgatgtgtggacaacacatctaattgtattttgacagcaatgggtgctaagaagaaaaaagacatggtaat
tgagataacaatggagagaagtcagagatggccttttctgaggagtgacatgctcagaatcgaataaaaagag
caggagcagcctttcaaggtgggtgcaaggatattcccggaagaaagaataagtggtgcaaaagccccagt
gggaaaaagcttggcaagctgaggtggttaaaagagctatctgtactaatgtctcttgtactgagttagcaag
gacaagagtgccaggggggtggaactggagagatagcagggaccacatctcacaggatctcgccagcctttt
aaggtatttggattttattgttagtgcaacaaggagccactggagagttttaaactagtggtgtgacctta
ctgtttcagaaagaacactggctactatgggaagaaaggacagtaggaagactaatagataatggtggattg
aactaagatggttagcaacagatagggagctatggtaatgttcagtatccactttggagatagatccagcagg
acttgctgacagactgcatgtaagggctgagggaaagagagctagcaaggtgacttctagtttgtgacctga
actaggtagatggttgtgttaaaaaatgcagtggttatttcagtcctgaagctcaggggcaatgatggaaaaa
agataaatgtgtgcaagttggctgggctggtggtgcacgcctataatcacagcactttgggagggccgagac
gggtggatcacttgaggtcagggcttggcagaccagcctaactgggtgaaacccatctctaataatacaaaa
aaaattagctgggcatggtggtgcatgctgtaatccaagctacttgggagggctgagacaagagaatcgctt
gcacctgggagggcgaggtttcagtgagctgagatcgaccattgcactccagcctgggcaacaagagcgaa
actccatctaaaaaaaagtgtgcaagttatttggtagctatatatttaagcctcgggactgaatgagatta
cctagagagaatgaagatagggaaacttcttgagctgagatcctataacatttagaagagaagctagccag
ggaaatttagaggggactggccagtgagagagagtaaaatctgaagagtataatgccatagtgtgaagctcag
agaattattgtttcaggaaggcagaagttgtcaattatgtccaacattgctgagaagatgagtaaggaatgt
aaatgagtaaaaggcaggaatggctattgaatttgttaagacagaggtccttgggtgaccttcttaagagtcat
ttcagcatgacaaggatgcagacctaaactggaatagattttcagggaaagactgagaggcaaggaaggggaga
ctgcaaatatatgcaaatctatatattttgcttgaagggaaacagaaaacaggacaacagctggtttacaga
acttttatttggattctttccttgtctgttaataatataggtgcacctgggcatcatcgtgcccttctctt
ttgattctgctgttttgccttttaactctacagggcatttcttctggcaggagccttgtacacactctgggtgg
ttcatgaccaccagggacttgccctttcaaggcagaagcctgtgattcaaaaacttccattccttccaatcg
actattataaaaactcattttgtctagtgcctactcctttaccttaggatcctgggccaactgtcactaaa
gaaaacaaaccatttctcctcctgggggaatgtgcttctttcaagaaggtgcttagagaatagacattcta
gggtgacagcagttatgaaagaggtgtctctaatagaacctaacacttcttccaacatactccaggttcacct
cctttgaagacagaaagtactacaacttcagttgccaccctagtccaggttcaggttactcagtaggtt
aagagatggtgctgagaatgaagtcagagtttgagtcacaagtgccatttgccttctccttgggacttct
gtggcaggggtgtggaacaaatcagattgggtgcatatctgttctccttctgaaaatacaattcaaatacc
ttcatagtctgtgacagcaacttataaccaacagacaatgggtattttttattttatttgatggcactcccc
atagccagaacttctactgagtcattcatctgtattccaggaatactttccaaatggagtagtgaggcact
tccaatacagttatgtaatctgttaacagtggtcctcaaatgacatgtgaaggaatgatgtacttcttttt
cagactcacagatctcatatgttaattcagtcacaagtagcttagggcttatcacatgccaggccacaggtaa
tatgacaataatgaaaacaaaaccccttgctgtcatagaattcggtttctagtgggaagagacatacattaa
caaatcatatataagtaaaacattaattatgggtgataacagaaaaagaaagcaggggagagggatagtgtct

tagttcaggttgctgtaacaagagtacatagacaggggtggcttamacaacaaacattttctttctcacagat
ctagtggctgggaagtctaagaccaggggtgccagcatgaagaagttctgggaagggttacattctgggttc
ctcacagtgagggctcttcacgcccttataaatgcactaattctattaccacatgacctcatctaaaacta
actcacaaggccaaacctcctatgttgagattaaggcttcaacataggaattttgggggggacaaaaacatt
cagtctctagcagataggggccactgtgggaggggattacaatcttaaatagggttgtcaggggaaggcttca
tgaggggcaaagacctgaggagcatggaaatgatccacgcagatgcatgtctggggaagagcattccaggca
gaggcagtggaatggaccctcaggttaagagcatgtttagcaagttcagagaaaagcctcttaggtcaggtc
tgagcgcagagtggaaaatgtaacaggacagatcatgtaaagacttgtagaccacagtttagaggcttgga
ataactctgaggatgacatgaaggtattgggaggggtgttgagcagaggagtgatatgatctgggtttatat
tttagattattctgactgctatgttgaaaatagggttgagagcagaggggaagggcagcacaccagttaga
ccacttcccaaaaatccagatgatgatgggcttgaaaccaggagagttgctatagggatgtgaagcaatcaa
ttctagatattttgggacctgctgacaggtacataatgggtgtatgagagaaaaggcagtggtgagaatgactct
gaagtctgtggcctgagcaaacagaaacctggagttgccaatagctgagataaagtttctgggttaaagatc
aagagctcagttttggacaagaggggtgtgagattttctaatagacattcaggtggaggtgtggggaacacac
ctggatgtgatgagcctggaattcaaattcacatgtgtgaatttgaaaaaagattcagaaaagaggtgtgga
ctggaggagctatccacacacagatagcatttgaaagccacaagactgggggtgatgacgcaggagtaagta
taaatagaaaagagaagtggtccaaggactcagtccttacatcagaggtcaaagagacaagatggagaggtcaac
gagatatgaaggcttatgcagtgagtggttcggagaaacgaagaaacttccagaagaaaggagaggtcaac
tgctaacaagtcaagtgaatgagaactgggaaatgagttgacaagagcaatttggtcggcgtgatcagaga
atgtttaagagataatgtgaggggaagaaaatggaggcaggaagtatgaactactcttttgagtcataattgta
aaggaaaaaatgggggtgattttactcttttttttttaagataggaaaaaaagtatatatttatatgctg
aggagaagaatccagttgaaggggaataactgtggcttcaggagacaggaaaaactgcctaagccacatccttc
agcaagtgagtggttcaagtgagggtttgtctttgcttggggcacgggcggttcacccacagtgaaggaga
gaaggcagggcatgtgggttggtctgtagaaagcgggcagacggggtgggaacttatgaagactctctgat
tgttattttctctgttaagtaaaaggaaaggtcattggctaagaatgaagatggacgcgcaggtgtgggag
tcacgaagagaatgaaagaatgaataaaatagagaaaaatacaatgctattatataaacctatacatttaa
attactttcctcttcaaattagacccccctggaaggcaaatctctgtcatgttaagtttaggaaaagtcata
atctacttgagctaaagtatgaagaataaggtatatgatttaaacacataattctatttttgcttaaaccg
gggaagcctttaaggtaacgtagagtgtacccattaaaaaagataaaaaatgtcagcttccccctcccttaa
atatagtaaggctcaaaggtaattgatgttcataaccctaagggtactataacagctctgaagcaaaatattaa
atgtctctctgtagtatttttagttaacttttaattttatttggttaatttatgtgtagttaaccacatacttg
cctcaactatttcacataaatgtgagatgtcttaggcctctttgaactctgcatggaaaagaatgtcgagga
ggtatgtctgtttaagctgccactgctgctctcataaataatatctcacattgcataatagcccagataa
actgcagatcgcttagagcctgataaatgagaggaaactgtacagtttggtttccaaattgcccttgacag
ctaattcacaaatttactattttatagctgtaagattttaaaaataaaaaataacatagctgtaaatacagat
tatacagaagatagtctgtaaaaatgtaaatataatataagctgtaaatatacactatcggaagtaccaaactc
agtttttagtaggttagagcatcaagtaaaatggaaatgaaagatgtttcaaggaaaactccaggacatctgt
ggcagtactaaagaaaccttctctcaggttcccagcatgctattttattgatgtaacaacattttcaata
aagttggtaaattatcactatattactgtcttatagcaacatagcaagagcttttagagatcatataagtata
aaatgtgaattttaaaaaaacaatgaatatgcaggatttttattagggcaagcgtttccataaccataaatat
ttcttttaaaacaaataaatgtcccaagatctctgttagtgatccaaactaagtagaaattagtaaaattaat
tataaatgaacaatttcagcatataaaaccaacaagtcttttctagatttttaacactgtgacccaattgcat
tattttccaagttagaatgactaataatcaatgaatgtaaaagcaataattaatacagatgacattgtactt
ttccacagtaaaagaaataaacaatctaataattttataaatccatttttatatcacaaaaataacctttacta
agcaaattttttttaaaatctcaggaactatagacatgatgaaaagatagatatattttatataaataattcaaa
aatactgtcaggcaaggaaatgtaaaatccttattttgagttaaagaaaatgctataaagcaatgagttatca
aaatacagaagaggtattctaaaacaaatgaaaaaccaagatgatgaaatagtgacaactacttctaagtgtg
taacagatactgaaatgccaaggtgaaagtgaactgaattattttcttaaagcagtgagagaatatgtaacttt
caaaaatgcaagaagcacagcaaatctaactaatttaacttaccacctccttcaaataaaagcgagaacctctt
gggagaatttaagcaccattagcagacacatcttagagc

Figure 2A SEQ ID NO: 5

MACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKAMPSAGTLPWVQ
GIICNANNPCFRYPTEAPGVVGNFNKSIVARLFSDARLLLYSQKDTSMKDMRKVLRTLQOIKKSSSNL
KLQDFLVDNETFSGFLYHNLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLDQEVSE
LCGLPREKLAAAEVRLSNMDILKPIRLTNSTSPFPSKELAEATKTLHSLGTLAQELFSMRWSDMRQE
VMFLTNVNSSSSSTQIYQAVSRIVCGHPGGGLKIKSLNWEYEDNNYKALFGGNGTEEDAETFYDNSTTPYC
NDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVMAEVNKTQELAVFHDLEGMWEELSPKIWTF
MENSQEMDLVRMLLDSRDNDHFWEQQLDGLDWTAQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRT
ISRMECVNLNKLEPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMDIDNVERTNK
IKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQMPYPCYVDDIFLRVMSRS
MPLFMTLAWIYSVAVVIKGIYVEKEARKETMRIMGLDNSILWFSWFISLIPLLVSAAGLLVILKLGNNLL
PYSDPVSVFVFLSVFAVVTILQCFLISTLFSRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFASL
LSPVAFGFGCEYFALFEEQIGVQWDNLFESEVEDGDFNLTSVSMMLFDTFLYGVMTWYIEAVFPGQYGI
PRPWYFPCTKSYWFGESDEKSHPGSNQKRISIEICMEEEPHTLKLGVSIQNLVKVYRDGMKVAVDGLALNF
YEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQNLGVCQHNVLFDMLTVEEHI
WFYARLKGLSEKHVKAEMEQMALDVGLPSSKLKSKTSQSLSGMQRKLVSALAFVGGSKVILDEPTAGVDP
YSRRGIWELLLLKYRQRTIILSTHMDADVLGDRIAIISHGKLCCVGSSFLKKNQLGTGYLTLVKKDVE
SSLSSCRNSSSTVSYLKKEDESVSQSSDAGLGSDESHTLTIDVSAISNLIRKHVSEARLVEDIGHELTYV
LPYEAAKEGAFVELFHEIDRLSDLGISYGISETTLEEIFLKVAEESGVDAETSDGTLPARRNRRAFGDK
QSCLRPFTEDDAADPNDSIDPESRETDLLSGMDGKGSYQVKGWKL TQQQFVALLWKRLLIARRSRKGFFA
QIVLPAVFVCIALVFLVPPFGKYPSELELQPMWYNEQYTFVSNDAPEDTGTLELLNALT KDPGFGTRCME
GNPIPDTPCQAGEEWTAPVPQTIMDLFQNGNWTMNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQ
NTADILQDLTGRIISDYLVKTYVQIIAKSLKNKIWNNEFRYGGFSLGVSNTQALPPSQEVNDAIKQMKKHL
KLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVINNAILRANLQKGENPSHYGITAFN
HPLNLTQQLSEVALMTTSVDVLVSICVIFAMSFVPASFVFLIQERVSKAKHLQFISGVKPVYIWL SNFV
WDMCNYVVPATLVIIIFICFQQKSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNL
FIGINGSVATFVLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSW
DLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRILDGGGQNDILEIKEL
TKIYRRKRKPAVDRIKCVGIPPGEFCGLLGVNAGKSSTFKMLTGDTTVTRGDAFLNKNLSILSNIHEVHQN
GYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKVGWAIKRLGLVKYGEKYAGNYSGGNKRKLSTAMAL
IGGPPVFLDEPTTGMDPKARRFLWNCALSVVKEGRSVLTSMSMEECEALCTRMAIMVNGRFRCLGSVQH
LKNRFGDGYTIVVRIAGSNPDLKPVQDFFLAPPGSVLKEKHNMLQYQLPSSLSSLARIFSIQSQSKRL
HIEDYSVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV*

Figure 2B SEQ ID NO: 6

GTCCCTGCTGTGAGCTCTGGCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGGGGTG
CTGGCTGAGGGAACATGGCTTGTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCA
CTTTTCAGAAGAAGACAAACATGTCAGCTGTTACTGGAAGTGGCCTGGCCTCTATTTATCT
TCCTGATCCTGATCTCTGTTTCGGCTGAGCTACCCACCCTATGAACAACATGAATGCCATT
TTCCAAATAAAGCCATGCCCTCTGCAGGAACACTTCCTTGGGTTTCAGGGGATTATCTGTA
ATGCCAACAACCCCTGTTTCCGTTACCCGACTCCTGGGGAGGCTCCCGGAGTTGTTGGAA
ACTTTAACAATCCATTGTGGCTCGCCTGTTCTCAGATGCTCGGAGGCTTCTTTTATACA
GCCAGAAAGACACCAGCATGAAGGACATGCGCAAAGTTCTGAGAACATTACAGCAGATCA
AGAAATCCAGCTCAAACCTGAAGCTTCAAGATTTCTTGGTGGACAATGAAACCTTCTCTG
GGTTCCTGTATCACAACTCTCTCTCCCAAAGTCTACTGTGGACAAGATGCTGAGGGCTG
ATGTCATTCTCCACAAGGTATTTTTGCAAGGCTACCAGTTACATTGACAAGTCTGTGCA
ATGGATCAAAATCAGAAGAGATGATTCAACTTGGTGACCAAGAAGTTTCTGAGCTTTGTG
GCCTACCAAGGGAGAACTGGCTGCAGCAGAGCGAGTACTTCGTTCCAACATGGACATCC
TGAAGCCAATCCTGAGAACACTAACTCTACATCTCCCTTCCCGAGCAAGGAGCTGGCTG
AAGCCACAAAAACATTGCTGCATAGTCTTGGGACTCTGGCCCAGGAGCTGTTTCAGCATGA
GAAGCTGGAGTGACATGCGACAGGAGGTGATGTTTCTGACCAATGTGAACAGCTCCAGCT
CCTCCACCCAAATCTACCAGGCTGTGTCTCGTATTGTCTGCGGGCATCCCGAGGGAGGGG
GGCTGAAGATCAAGTCTCTCAACTGGTATGAGGACAACAACCTACAAAGCCCTCTTTGGAG
GCAATGGCACTGAGGAAGATGCTGAAACCTTCTATGACAACCTCTACAACCTCTTACTGCA
ATGATTTGATGAAGAATTTGGAGTCTAGTCCTCTTTCCCGCATTATCTGGAAAGCTCTGA
AGCCGCTGCTCGTTGGGAAGATCCTGTATACACCTGACACTCCAGCCACAAGGCAGGTCA
TGGCTGAGGTGAACAAGACCTTCCAGGAACTGGCTGTGTTCCATGATCTGGAAGGCATGT
GGGAGGAACTCAGCCCCAAGATCTGGACCTTCATGGAGAACAGCCAAGAAATGGACCTTG
TCCGGATGCTGTTGGACAGCAGGGACAATGACCACTTTTGGGAACAGCAGTTGGATGGCT
TAGATTGGACAGCCCCAAGACATCGTGGCGTTTTTTGGCCAAGCACCCAGAGGATGTCCAGT
CCAGTAATGGTTCTGTGTACACCTGGAGAGAAGCTTTCAACGAGACTAACCAGGCAATCC
GGACCATATCTCGCTTCATGGAGTGTGTCAACCTGAACAAGCTAGAACCCATAGCAACAG
AAGTCTGGCTCATCAACAAGTCCATGGAGCTGCTGGATGAGAGGAAGTTCTGGGCTGGTA
TTGTGTTCACTGGAATTACTCCAGGCAGCATTGAGCTGCCCCATCATGTCAAGTACAAGA
TCCGAATGGACATTGACAATGTGGAGAGGACAAATAAAATCAAGGATGGGTACTGGGACC
CTGGTCCCTCGAGCTGACCCCTTTGAGGACATGCGGTACGTCTGGGGGGGCTTCGCCTACT
TGCAGGATGTGGTGGAGCAGGCAATCATCAGGGTGCTGACGGGCACCGAGAAGAAAACCTG

GTGTCTATATGCAACAGATGCCCTATCCCTGTTACGTTGATGACATCTTTCTGCGGGTGA
 TGAGCCGGTCAATGCCCCCTCTTCATGACGCTGGCCTGGATTTACTCAGTGGCTGTGATCA
 TCAAGGGCATCGTGTATGAGAAGGAGGCACGGCTGAAAGAGACCATGCGGATCATGGGCC
 TGGACAACAGCATCCTCTGGTTTAGCTGGTTCATTAGTAGCCTCATTCCTCTTCTTGTGA
 GCGCTGGCCTGCTAGTGGTCATCCTGAAGTTAGGAAACCTGCTGCCCTACAGTGATCCCA
 GCGTGGTGTGTTGTCTTCTGTCCGTGTTTGCTGTGGTGACAATCCTGCAGTGCTTCCTGA
 TTAGCACACTCTTCTCCAGAGCCAACCTGGCAGCAGCCTGTGGGGGCATCATCTACTTCA
 CGCTGTACCTGCCCTACGTCCTGTGTGTGGCATGGCAGGACTACGTGGGCTTCACACTCA
 AGATCTTCGCTAGCCTGCTGTCTCCTGTGGCTTTTGGGTTTGGCTGTGAGTACTTTGCCC
 TTTTGTAGGAGCAGGGCATTGGAGTGCAGTGGGACAACCTGTTTGAGAGTCCTGTGGAGG
 AAGATGGCTTCAATCTCACCCTTCGGTCTCCATGATGCTGTTTGACACCTTCCTCTATG
 GGGTGATGACCTGGTACATTGAGGCTGTCTTCCAGGCCAGTACGGAATTCACAGGCCCT
 GGTATTTTCTTGCACCAAGTCCTACTGGTTTGGCGAGGAAAGTGATGAGAAGAGCCACC
 CTGGTTCCAACCAGAAGAGAATATCAGAAATCTGCATGGAGGAGGAACCCACCCACTTGA
 AGCTGGGCGTGTCCATTCAGAACCTGGTAAAAGTCTACCGAGATGGGATGAAGGTGGCTG
 TCGATGGCCTGGCACTGAATTTTTATGAGGGCCAGATCACCTCCTTCCTGGGCCACAATG
 GAGCGGGGAAGACGACCACCATGTCAATCCTGACCGGTTGTTCCCCCGACCTCGGGCA
 CCGCCTACATCCTGGGAAAAGACATTCGCTCTGAGATGAGCACCATCCGGCAGAACCTGG
 GGGTCTGTCCCCAGCATAACGTGCTGTTTGACATGCTGACTGTGCAAGAACACATCTGGT
 TCTATGCCCCGCTTGAAAGGGCTCTCTGAGAAGCACGTGAAGGCGGAGATGGAGCAGATGG
 CCCTGGATGTTGGTTTGCCATCAAGCAAGCTGAAAAGCAAAACAAGCCAGCTGTGAGGTG
 GAATGCAGAGAAAGCTATCTGTGGCCTTGGCCTTTGTGCGGGGATCTAAGGTTGTATTCT
 TGGATGAACCCACAGCTGGTGTGGACCTTACTCCCGCAGGGGAATATGGGAGCTGCTGC
 TGAAATACCGACAAGGCCGCACCATATTCTCTCTACACACCACATGGATGAAGCGGACG
 TCCTGGGGGACAGGATTGCCATCATCTCCCATGGGAAGCTGTGCTGTGTGGGCTCCTCCC
 TGTTTCTGAAGAACCAGCTGGGAACAGGCTACTACCTGACCTTGGTCAAGAAAGATGTGG
 AATCCTCCCTCAGTTCTGCAGAAACAGTAGTAGCACTGTGTCTACCTGAAAAAGGAGG
 ACAGTGTTTCTCAGAGCAGTTCTGATGCTGGCCTGGGCAGCGACCATGAGAGTGACACGC
 TGACCATCGATGTCTCTGCTATCTCCAACCTCATCAGGAAGCATGTGTCTGAAGCCCGGC
 TGGTGGAAGACATAGGGCATGAGCTGACCTATGTGCTGCCATATGAAGCTGCTAAGGAGG
 GAGCCTTTGTGGAACCTTTTCATGAGATTGATGACCGGCTCTCAGACCTGGGCATTTCTA
 GTTATGGCATCTCAGAGACGACCCTGGAAGAAATATTCCTCAAGGTGGCCGAAGAGAGTG

GGGTGGATGCTGAGACCTCAGATGGTACCTTGCCAGCAAGACGAAACAGGCGGGCCTTCG
GGGACAAGCAGAGCTGTCTTCGCCCCGTTCACTGAAGATGATGCTGCTGATCCAAATGATT
CTGACATAGACCCAGAATCCAGAGAGACAGACTTGCTCAGTGGGATGGATGGCAAAGGGT
CCTACCAGGTGAAAGGCTGGAACTTACACAGCAACAGTTTGTGGCCCTTTTGTGGAAGA
GACTGCTAATTGCCAGACGGAGTCGGAAAGGATTTTTTGCTCAGATTGTCTTGCCAGCTG
TGTTTGTCTGCATTGCCCTTGTGTTTCAGCCTGATCGTGCCACCCTTTGGCAAGTACCCCA
GCCTGGAACCTCAGCCCTGGATGTACAACGAACAGTACACATTTGTCAGCAATGATGCTC
CTGAGGACACGGGAACCCTGGAACCTTAAACGCCCTCACCAAAGACCCTGGCTTCGGGA
CCCGCTGTATGGAAGGAAACCAATCCCAGACACGCCCTGCCAGGCAGGGGAGGAAGAGT
GGACCACTGCCCCAGTTCCCCAGACCATCATGGACCTCTTCCAGAATGGGAACCTGGACAA
TGCAGAACCCCTTCACCTGCATGCCAGTGTAGCAGCGACAAAATCAAGAAGATGCTGCCTG
TGTGTCCCCCAGGGGCAGGGGGCTGCCTCCTCCACAAAGAAAACAAAACACTGCAGATA
TCCTTCAGGACCTGACAGGAAGAAACATTTTCGGATTATCTGGTGAAGACGTATGTGCAGA
TCATAGCCAAAAGCTTAAAGAACAAGATCTGGGTGAATGAGTTTAGGTATGGCGGCTTTT
CCCTGGGTGTCTAGTAATACTCAAGCACTTCCTCCGAGTCAAGAAGTTAATGATGCCATCA
AACAAATGAAGAAACACCTAAAGCTGGCCAAGGACAGTTCTGCAGATCGATTTCTCAACA
GCTTGGGAAGATTTATGACAGGACTGGACACCAGAAATAATGTCAAGGTGTGGTTCAATA
ACAAGGGCTGGCATGCAATCAGCTCTTTCTGAATGTCAACAATGCCATTCTCCGGG
CCAACCTGCAAAAGGGAGAGAACCCTAGCCATTATGGAATTACTGCTTTCAATCATCCCC
TGAATCTCACCAAGCAGCAGCTCTCAGAGGTGGCTCTGATGACCACATCAGTGGATGTCC
TTGTGTCCATCTGTGTCTATCTTTGCAATGTCTTCGTCCCAGCCAGCTTTGTCTGATTCC
TGATCCAGGAGCGGGTCAGCAAAGCAAAACACCTGCAGTTCATCAGTGGAGTGAAGCCTG
TCATCTACTGGCTCTCTAATTTTGTCTGGGATATGTGCAATTACGTTGTCCCTGCCACAC
TGGTCATTATCATCTTCATCTGCTTCCAGCAGAAGTCCTATGTGTCTCCACCAATCTGC
CTGTGCTAGCCCTTCTACTTTTGTCTGATGGGTGGTCAATCACACCTCTCATGTACCCAG
CCTCCTTTGTGTTCAAGATCCCCAGCACAGCCTATGTGGTGCTCACCAGCGTGAACCTCT
TCATTGGCATTAAATGGCAGCGTGGCCACCTTTGTGCTGGAGCTGTTACCGACAATAAGC
TGAATAATATCAATGATATCCTGAAGTCCGTGTTCTTGATCTTCCCACATTTTTGCCTGG
GACGAGGGCTCATCGACATGGTGAAAAACCAGGCAATGGCTGATGCCCTGGAAAGGTTTG
GGGAGAATCGCTTTGTGTCACCATTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCCA
TGGCCGTGGAAGGGGTGGTGTCTTCTCCTCATTACTGTTCTGATCCAGTACAGATTCTTCA
TCAGGCCCAGACCTGTAAATGCAAAGCTATCTCCTCTGAATGATGAAGATGAAGATGTGA
GGCGGGAAAGACAGAGAATTCTTGATGGTGGAGGCCAGAATGACATCTTAGAAATCAAGG
AGTTGACGAAGATATATAGAAGGAAGCGGAAGCCTGCTGTTGACAGGATTTGCGTGGGCA

TTCTCCTGGTGAGTGCTTTGGGCTCCTGGGAGTTAATGGGGCTGGAAAATCATCAACTT
 TCAAGATGTTAACAGGAGATACCACTGTTACCAGAGGAGATGCTTTCCTTAACAAAAATA
 GTATCTTATCAAACATCCATGAAGTACATCAGAACATGGGCTACTGCCCTCAGTTTGATG
 CCATCACAGAGCTGTTGACTGGGAGAGAACACGTGGAGTTCTTTGCCCTTTTGAGAGGAG
 TCCCAGAGAAAAGAAGTTGGCAAGGTTGGTGAGTGGGCGATTTCGAAACTGGGCCTCGTGA
 AGTATGGAGAAAAATATGCTGGTAACTATAGTGGAGGCAACAAACGCAAGCTCTCTACAG
 CCATGGCTTTGATCGGCGGGCCTCCTGTGGTGTCTTCTGGATGAACCCACCACAGGCATGG
 ATCCCAAAGCCCGCGGTTCTTGTGGAATTGTGCCCTAAGTGTGTCAAGGAGGGGAGAT
 CAGTAGTGCTTACATCTCATAGTATGGAAGAATGTGAAGCTCTTGCCTAGGATGGCAA
 TCATGGTCAATGGAAGGTTTCAAGTGCCTTGGCAGTGTCCAGCATCTAAAAAATAGGTTTG
 GAGATGGTTATACAATAGTTGTACGAATAGCAGGGTCCAACCCGACCTGAAGCCTGTCC
 AGGATTTCTTTGGACTTGCAATTTCTTGAAGTGTCTAAAAGAGAAACACCGGAACATGC
 TACAATACCAGCTTCCATCTTCATTATCTTCTCTGGCCAGGATATTCAGCATCCTCTCCC
 AGAGCAAAAAGCGACTCCACATAGAAGACTACTCTGTTTCTCAGACAACACTTGACCAAG
 TATTTGTGAACTTTGCCAAGGACCAAAGTGATGATGACCCTTAAAAGACCTCTCATTAC
 ACAAAAACCAGACAGTAGTGACGTTGCAGTTCTCACATCTTTTCTACAGGATGAGAAAAG
 TGAAAGAAAGCTATGTATGAAGAATCCTGTTTATACGGGGTGGCTGAAAGTAAAGAGGAA
 CTAGACTTTCTTTGCACCATGTGAAGTGTGTGGAGAAAAGAGCCAGAAGTTGATGTGG
 GAAGAAGTAAACTGGATACTGTACTGATACTATTCAATGCAATGCAATTCAATGCAATGA
 AAACAAAATTCATTACAGGGGCGAGTGCCTTTGTAGCCTATGTCTTGTATGGCTCTCAAG
 TGAAAGACTTGAAATTTAGTTTTTTACCTATACCTATGTGAAACTCTATTATGGAACCCAA
 TGGACATATGGGTTTGAACCTCACACTTTTTTTTTTTTTTTTTTGTTCCTGTGTATTCTCATT
 GGGGTTGCAACAATAATTCATCAAGTAATCATGGCCAGCGATTATTGATCAAAATCAAAA
 GGTAATGCACATCCTCATTCACTAAGCCATGCCATGCCAGGAGACTGGTTTCCCGGTGA
 CACATCCATTGCTGGCAATGAGTGTGCCAGAGTTATTAGTGCCAAGTTTTTTCAGAAAGTT
 TGAAGCACCATGGTGTGTCTGCTCACTTTTGTGAAAGCTGCTCTGCTCAGAGTCTATCA
 ACATTGAATATCAGTTGACAGAATGGTGCCATGCGTGGCTAACATCCTGCTTTGATTCCC
 TCTGATAAGCTGTTCTGGTGGCAGTAACATGCAACAAAAATGTGGGTGTCTCCAGGCACG
 GGAAACTTGGTTCCATTGTTATATTGTCCTATGCTTCGAGCCATGGGTCTACAGGGTCAT
 CCTTATGAGACTCTTAAATATACTTAGATCCTGGTAAGAGGCAAAGAATCAACAGCCAAA
 CTGCTGGGGCTGCAACTGCTGAAGCCAGGGCATGGGATTAAAGAGATTGTGCGTTCAAAC
 CTAGGGAAGCCTGTGCCCATTTGTCCTGACTGTCTGCTAACATGGTACACTGCATCTCAA
 GATGTTTATCTGACACAAGTGTATTATTTCTGGCTTTTGAATTAATCTAGAAAATGAAA

Figure 3

Promoter, 8797 bp

Distances numbered using first base of promoter as 1

Name	Pos. of 1st base in sense strand	Hit Site	% Match	Strand	SEQ ID No.
LXRE		Target: AGGTCA (NNNN)AGGTCA			7
DR4	-7531	AGAGCAGGTGCATCATTTGAGGTCA	88	sense	8
DR4	-5085	TTGAGCGGGTGATCACTTGAGGTCA	88	antisense	9
DR4	-4389	CAAGCGGGCAGATCACTTGAGGTCA	88	antisense	10
DR4	-1641	CAAGTGGGCAAGTCACTCAGGTCA	94	antisense	11
DR1		None			
PPAR		Target: NNNNNN (A) NN (T) TGACCT (N/NN) TGACCT			12
DR2	-7718	CTTTGA (A) GC (C) TGATCATATGACCT	88	antisense	13
DR2	-7521	AGGCTG (G) TC (T) CGAACTCCTGACCT	88	antisense	14
DR2	-5708	CTTAAT (T) GG (T) GGGCTGTTGACCT	91	antisense	15
DR2	-2894	CAGGAT (G) GC (G) TAACTCCTGACCT	88	antisense	16
DR2	-1649	AGGTTG (G) TT (T) CGAACTCCTGACCT	88	sense	17
DR2	-1140	TCAAGG (T) AG (G) AGACCTTGTGGCCT	88	sense	18
DR1		None			

Name	Pos. of 1st base in sense strand	Hit Site	% Match	Strand
SREBP		Target: ATCACCCAC		
	-8523	GAGATGTGCTATGACCCAC	90	antisense
	-3651	GTGAGCCCGATCACACCAC	90	antisense
	-7747	TCCATCCATCCACACCCAC	80	antisense
	-5485	CCCTTTTATTAAACACCTCAC	80	antisense
	-5248	GTAAGCCAAAGATCATGCCAC	80	antisense
	-5073	ACCTCAAAGTGATCACCGGCC	80	sense
	-2252	GGCTCAAAGCGATCCTCCAC	80	antisense
	-2209	CCATGATTGGATCACTGCAC	80	sense
	-1794	GTGAGTCGAGATCATGCCAC	80	antisense
	-519	TGCTTTTGTGTTTCCCCAC	80	antisense
	-478	CCGCCTTCCCTCACCCGAG	80	sense
	-158	ACCTCCACCCCCACCCAC	80	sense

Page 2 of 6

Exon 1, 303 bp

Distances numbered using start of Exon 1 as 1

Name	Pos. of 1st base in +	Hit Site	% Match	Strand	Q ID NO.
LXRE		Target: AGGTCA (NNNN) AGGTCA			7
DR4	4	CCGAGCGCAGAGGTTACTATCGGTCA	92	antisense	38
DR1		None			

PPAR					
DR2		None			
DR1		None			
SREBP		None			
ROR		None			

5' Intron 1, 930 bp

Positions numbered using the first position in intron 1 as + 1

Name	Pos. of 1st base in +	Hit Site	Matc	Strand	SEQ ID N
LXRE		Target: AGGTCA (NNNN)AGGTCA			7
DR4	458	GCCAAATCCCAGGTCAACAGACCA	88	antisense	39
DR1		None			
PPAR		Target: NNNNNN (A) NN (T) TGACCT (N/NN)TGACCT			12
DR2		None			
DR1		None			
SREBP		Target: ATACCCCAC			19
	326	GGACCTGCAGCTCTCCCCAC	80	antisense	40
ROR		Target: (w){0,8}WRGGTCA			32
	17	AACGCCCAAGTAAGTCA	94	antisense	41
	161	GAGCTCGTACTAGGACA	94	antisense	42
	181	GCAGAGTCCTGGGGTCA	94	antisense	43
	181	CGCAGAGTCCTGGGGTCA	94	antisense	44
	478	AGCCAATCCAGGTCA	94	antisense	45
	559	ACGGACCGTTTGGGACA	94	antisense	46
	559	CACGGACCGTTTGGGACA	94	antisense	47
	559	CCACGGACCGTTTGGGACA	94	antisense	48
	589	ACTAGAGCCCTTGGGTCT	94	sense	49
	590	CTAGAGCCCTTGGGTCT	94	sense	50
	612	CCCTACCCCTCAGGTCA	94	antisense	51
	612	TCCCTACCCCTCAGGTCA	94	antisense	52
	668	GSTCTGCGCCAGGACA	94	antisense	53
	864	TTTTAGTGAGANGGTTA	94	sense	54

3' end of Intron 1, 12504 bp

Positions numbered using the first base 5 to the start of Exon 2 as -1

Name	Pos. of 1st base in +	Hit Site	Mat	Strand	
LXRE		Target: AGGTCA (NNNN)AGGTCA			7
DR4	-7188	TGAGGCAGGTAGATCACTTGAGGTCA	93	sense	55
DR4	-11050	CGAGGCTGGCGGATCACTTGAGGTCA	86	sense	56
DR4	-7670	AAGCCTAACAAAGGTTACTGAAGGCCA	86	antisense	57
DR4	-4696	AGAGTGGCGGATCACTTGAGGTCA	86	antisense	58
DR1		None			
PPAR		Target: NNNNNN (A) NN (T) TGACCT (N/NN) TGACCT			12
DR1		None			
DR2	-10281	CTCGAT (T) TC (C) TGACCTCGTGATCC	86	antisense	59
DR2	-5996	CAAAAC (A) TT (G) TGCCCTTTTGAACCT	86	antisense	60
DR2	-932	GCGCTA (G) GG (T) TGTCCTCATTAACCT	86	sense	61
DR2	-597	CTCGAT (T) TC (T) TGACCTCGTGATCC	86	sense	62
SREBP		Target: ATCACCCAC			19
	-7009	GTGAGCTCAGATCACACCAC	90	sense	63
	-11869	TTCAAGGATGATCACACAT	80	antisense	64
	-11616	GGCTCAAGTGATCCTCCAC	80	antisense	65
	-10100	GTGAGCCGAGATCGGCCAC	80	sense	66
	-8584	GTGAGTTATGATCATGCCAC	80	antisense	67
	-5591	CCACTGTTTGAACAACCCAC	80	sense	68
	-4684	ACCTCAGGTGATCGGCCAC	80	sense	69
	-4128	AAATGTGACAAATCTCCACAC	80	antisense	70
	-2524	AAATAGATATACCTCCC	80	antisense	71
	-1577	CCTTTTATCTACCAACCAAC	80	antisense	72

Figure 4

Mutations:																									
Exon	Numbering based on Pullinger et al., 2000	Numbering based on AJ012376.1	Sequence difference/context	SEQ ID NO	Change in codon or non-coding	Coding Nucl. Change	AA change	Genotypes		ABE	MGA	ALA	RLA	TDW	TDS	NL16	NL27	QC11	QC12	CZ01	JP01	NL16-212	NL20	BC11	BAC
Exon 13	2154	1699	TTGAGGACATGCGGTACGTCGTGG TTGAGGACATGTGTACGTCGTGG	84 85	CGG-TGG	C2154T	R567W	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 13	2185	1730	GCCTACTTGCAGGATGTGGTG GCCTACTTGCAGGATGTGGTG	86 87	CAG-CGG	A2185G	Q597R	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 14	2472-2474	2017-2019	CCCTCAATTCCTCTTCTTGTGAGCG CCTCAATTCCTCTTCTTGTGAGCG	88 89	Deletion of L (CTT)	DeltaCT2472-2474	DeltaL693	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 18	3120	2665	AAAAGTCTACCGAGATGGGAT AAAAGTCTACTGAGATGGGAT	90 91	CGA-TGA	C3120T	R000Stop	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 18	3181	2726	GGCCAGATCAGCTCTCTCTCTG GGCCAGATCATCTCTCTCTCTG	92 93	AGC-ATC	C3181T	T929I	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 22	3667	3212	ACACACCACATGGATGAAGCG ACACACCACAGGATGAAGCG	94 95	ATG-ACG	T3667C	M1091T	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Intron 24	(+) 1	(+) 1	CCTGGAAGAAGTAAGTTAAGT CCTGGAAGAAGTAAGTTAAGT	96 97	NA (G to C in splice junct.)	IVS24 + TG to C	Not applicable	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 30	4824	4369	GCTGCCGTGTGTGCCGCCAGG GCTGCCGTGTGTGCCGCCAGG	98 99	TGT-CGT	T4824C	C1477R	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 35	5277-5278	4822-23	TAGCCATTATGGAATTACTGCT TAGCCATTATCAATTACTGCT	100 101	NA (GG to C)	GG5277-5278C	Frameshift1628	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 41	6073-6078	5618-5623	GATGAAGATGAAGATGTAGGGCGGGA GATGAAGATGTAGGGCGGGA	102 103	Deletion of E.D (AATATG)	DeltaAATATG6073-8	Delta(E.D)1893-4	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 48	6825	6370	AATAGTTGTACGAATAGCAGG AATAGTTGTATGAATAGCAGG	104 105	CGA-TGA	C6825T	R2144Stop	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 48	6825	6370	ATAGTTGTACGAATAGCAGG ATAGTTGTATGAATAGCAGG	106 107	NA (Deletion of C)	DeltaC6825	2145Stop	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 48	6844	6369	GGGTCCAAACCCGGACCTGAA GGGTCCAAACCTGGACCTGAA	108 109	CCG-CTG	C6844T	P2150L	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt
Exon 48	6852-6854	6497-6499	CATTATCTCTCTGGCCAGGA CATTATCTTTTGGCCAGGA	110 111	NA (CTC to TT)	CTC6852-6854TT	2203Stop	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt	wt

Coding Polymorphisms (cSNP):									
Exon	Numbering based on Pullinger et al., 2000	Numbering based on AJ012376.1	Sequence	Change in codon	Coding Nucl. Change	AA change	Genotypes	ABE	MLA
Exon 0/5'UT	69	NA	Wildtype Variant	NA	C69T	Not applicable	C/C	C/C	C/C
Exon 0/5'UT	127	NA	Wildtype Variant	NA	C127G	Not applicable	C/C	C/C	C/C
Exon 1/5'UT	319	(-)256	Wildtype Polymorph	NA	InsG319	Not applicable	wt	wt	wt
Exon 1/5'UT	378	(-)198	Wildtype Polymorph	NA	G378C	Not applicable	wt	wt	wt
Exon 5	869	414	Wildtype Polymorph	CTG-CTA	G869A	No aa change	G/G	G/G	G/G
Exon 6	1051	596	Wildtype Polymorph	AGG-AAG	G1051A	R219K	wt	wt	wt
Exon 8	1331	876	Wildtype Polymorph	CCC-CCT	C1331T	No aa change	wt	wt	wt
Exon 8	1343	888	Wildtype Polymorph	GGG-GGA	G1343A	No aa change	wt	wt	wt
Exon 10	1591	1136	Wildtype Polymorph	GTC-GCG	T1591C	V399A	wt	wt	wt
Exon 15	2706	2251	Wildtype Polymorph	GTG-ATG	G2706A	V771M	wt	wt	wt
Exon 15	2715	2260	Wildtype Polymorph	ACA-CCA	A2715C	T774P	wt	wt	wt
Exon 15	2723	2268	Wildtype Polymorph	AAG-AAC	G2723C	K778N	wt	wt	wt
Exon 16	2868	2413	Wildtype Variant	GTC-ATC	G2868A	V825I	wt	wt	wt
Exon 17	3044	2589	Wildtype Polymorph	ATA-ATG	A3044G	I863M	wt	wt	wt
Exon 21	3554	3099	Wildtype Polymorph	GTT-GTG	T3554G	No aa change	wt	wt	wt
Exon 23	3911	3456	Wildtype Polymorph	GAG-GAC	G3911C	E1172D	wt	wt	wt
Exon 34	5155	4700	Wildtype Polymorph	AGA-AAA	G5155A	R1587K	wt	wt	wt
Exon 37	5587	5132	Wildtype	TCC-TGC	C5587G	S1731C	wt	wt	wt

Non-coding Polymorphisms (SNP):									
Intron	Numbering based on Pullinger et al., 2000	Numbering based on AJ012376.1	Sequence	Change in non-coding	Coding Nucl. Change	AA change	Genotypes	ABE	MGA
Promoter	(-) 191	NA	CGGGGAGGACGACGACG	G to C	Not applicable	Not applicable	C/C	C/C	C/C
Promoter	(-) 17	NA	GGCGGGAAGCGGCGG	C to G	Not applicable	Not applicable	C/C	C/C	C/C
Intron 0	(+) 1163	(-) 1435	AGTATCCCTGTTCACGAGAA	CCCT insertion	Not applicable	Not applicable	no ins	no ins	no ins
Intron 0	(-) 1095	(-) 1367	GTGACACCCACGAGTAGGG	A to G	Not applicable	Not applicable	A/A	A/A	A/A
Intron 0	(-) 1027	(-) 1299	TATGTGCTGACCATGGAGCTTGT	G to A	Not applicable	Not applicable	A/A	A/A	A/A
Intron 0	(+) 720	(-) 992	CCTCGCTGCGAGTTTCAGCGATT	G to A	Not applicable	Not applicable	A/A	A/A	A/A
Intron 0	(+) 481	(-) 733	GAAATTAAGTAGTGTAAAGGAG	A to C	Not applicable	Not applicable	A/A	A/A	A/A
Intron 0	(+) 362	(+) 634	CATTTCCTTAGAAGAGAGGT	A to G	Not applicable	Not applicable	A/A	A/A	A/A
Intron 7	(+) 2383	(+) 2383	TTTAAAGGGGKTGATTAGGA	Frequencies unknown; G and T	Not applicable	Not applicable			
Intron 7	(+) 3035	(+) 3035	GAAGAAATTTTCTTTTGTATT	Frequencies unknown; G and T	Not applicable	Not applicable			
Intron 7	(-) 15	(-) 15	TCTGTCCCAATCCCTGACG	A insertion	Not applicable	Not applicable	wt	wt	wt
Intron 9	(-) 42	(-) 42	AGGAGCCAAACGCTCATTTGT	G insertion	Not applicable	Not applicable	wt	wt	wt
Intron 13	(+) 24	(+) 24	AAGCCACTGTTTAAACAGT	T to A	Not applicable	Not applicable	wt	wt	wt
Intron 13	(-) 83	(-) 83	GCTCCCTGTAGGATGACAGGCTC	C to T	Not applicable	Not applicable	wt	wt	wt
Intron 15	(-) 4-5	(-) 4-5	TTGCGCTGTTCTCAGAGGOC	CA deletion	Not applicable	Not applicable	wt	wt	wt
Intron 17	(+) 2000	(+) 2000	GCGCAGTGCSCGTGTCTCTTA	Frequencies unknown; C and G	Not applicable	Not applicable			
Intron 21	(+) 118	(+) 118	CTCTCTGTTAKACAGAGAAGA	Frequencies unknown; G and T	Not applicable	Not applicable			

Errors in public sequence (differences between all samples and Genbank entry AJ012376.1):																			
Exon/Intron	Numbering based on Pullinger et al., 2000	Numbering based on AJ012376.1	Sequence difference/context in cDNA	Change in codon or non-coding	Coding Nucl. Change	AA change	Genotypes		ABE	MGA	ALA	RLA	TDW	TDS	NL16	NL27	QC11	QC12	C201
Exon 2	16, 18	Public seq	TGTCAGCTGTTACTGGAAAGTGG	T to C, A to G		No aa change													
		Correct se	TGTCAGCTGCTGCTGGAAGTGG																
Exon 7	705	Public seq	AGGAGCTGGCGGAAGCCACAA	C to T		No aa change													
		Correct se	AGGAGCTGGCTGAAGCCACAA																
Exon 33	4604	Public seq	AATGATGCCGCAACAAATG	ACC-ATC		Thr to Ile													
		Correct se	AATGATGCCATCAACAAATG																
Exon 35	4883	Public seq	GAGGTGGCTCCGATGACCACA	CCG-CTG		Pro to Leu													
		Correct se	GAGGTGGCTCTGTGATGACCACA																
Exon 43	5861	Public seq	TTCCTTAACAGAAATAGTATC	AGA-AAA		Arg to Lys													
		Correct se	TTCCTTAACAAAAATAGTATC																
Exon 48	6443	Public seq	GGAAGTGTTCACAAAGAGAAA	CCA-CTA		Pro to Leu													
		Correct se	GGAAGTGTTCATAAAGAGAAA																
Exon 49	6765	Public seq	AGTAAAGAGGGAGCTAGACTTT	G to A		No aa change													
		Correct se	AGTAAAGAGGAAGCTAGACTTT																

Figure 5A

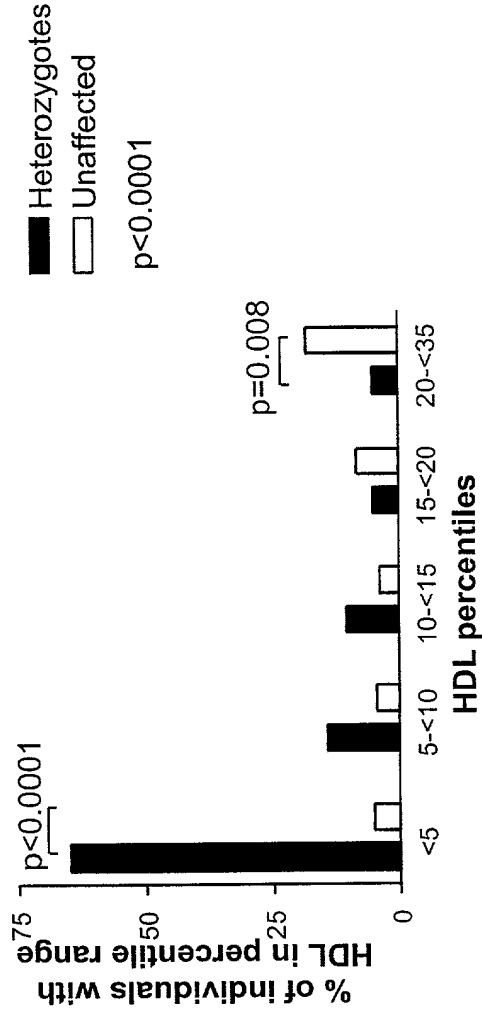


Figure 5B

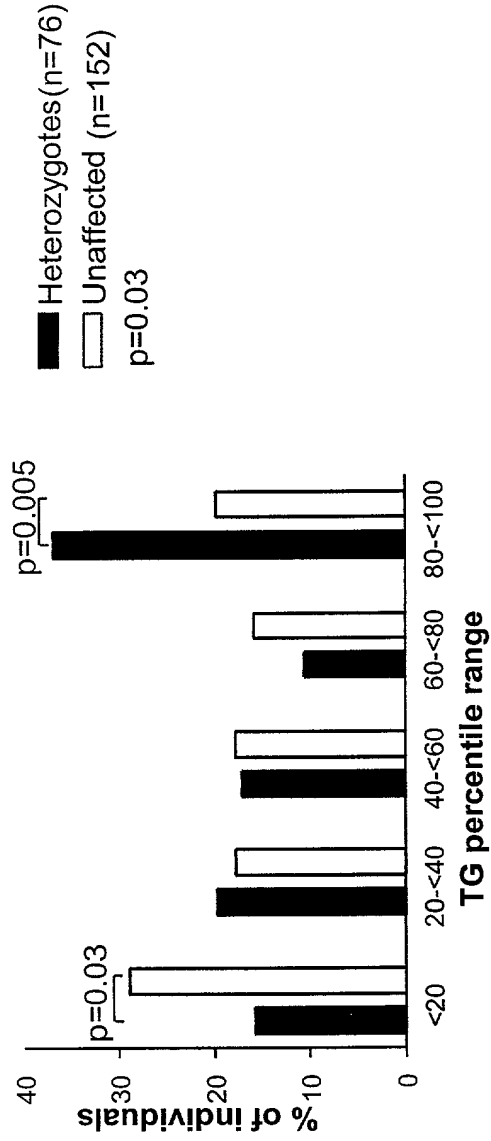


Figure 6

	TD Patients	Heterozygotes	Unaffected family members	P-value heterozygotes vs. unaffected	P-value TD patients vs. unaffected
number	5	77 ^A	156 ^A		
age (yrs) range	43.4±9.0 31-56	42.5±19.6 5-81	39.9±21.0 4-86	0.35	0.71
m/f	3/2	33/44	82/74	0.16	0.74
TC (mmol/L)	2.34±1.03	4.52±1.12	4.71±1.07	0.23	<0.0001
TG (mmol/L)	1.95±0.97	1.66±1.59	1.20±1.03	0.03	0.11
HDL (mmol/L)	0.08±0.05	0.74±0.24	1.31±0.35	<0.0001	<0.0001
LDL (mmol/L)	1.37±1.02	3.03±0.99	2.84±0.87	0.171	0.0003
ApoA-I (g/L)	0.03±0.04 (3)	0.92±0.32 (61)	1.43±0.26 (55)	<0.0001	<0.0001
ApoA-II (g/L)	0.10±0.08 (2)	0.35±0.08 (46)	0.39±0.08 (43)	0.01	<0.0001
ApoB (g/L)	0.89±0.53 (2)	0.93±0.25 (52)	0.94±0.33 (42)	0.88	0.84
CHD ≥ 20 yrs	20% (1/5)	12.9% (8/62)	4.1% (5/122)	0.03	0.10
Odds Ratio (95% CI)				3.47 (1.08-11.09)	5.85 (0.55-62.4)
Age of onset	38	48.9±8.6	60.4±12.8	0.08	

^A For TC, TG, LDL n=76 for heterozygotes, 153 for unaffected family members

Figure 7

Individual	Mutation	exon	disease (age of onset)
TD proband			
TD1	C1477R, ivs24+1G-->C 30, intron 24		CHD (38)
ABC1 heterozygotes			
TD4-201	unidentified	-	MI (<58)
FHA5-215	M1091T	22	MI (61)
FHA5-303	M1091T	22	CHD (<45)
TD1-363	C1477R	30	MI (51)
FHA3-301	Del(E,D) 1893,94	41	PVD (<54)
FHA3-305	Del(E,D) 1893,94	41	CHD (44)
FHA6-201	P2150L	48	CVA (36), fatal MI (58)
FHA2-301	R2144X	48	CAD (42), PTCA (47), femoral angioplasty (48), CABG (<50)
Unaffected family members			
FHA5-212	none	-	AP (62)
TD3-109	none	-	TIA (80)
FHA2-315	none	-	MI (51)
TD1-205	none	-	MI (62)
TD1-216	none	-	AP (47)

Figure 8

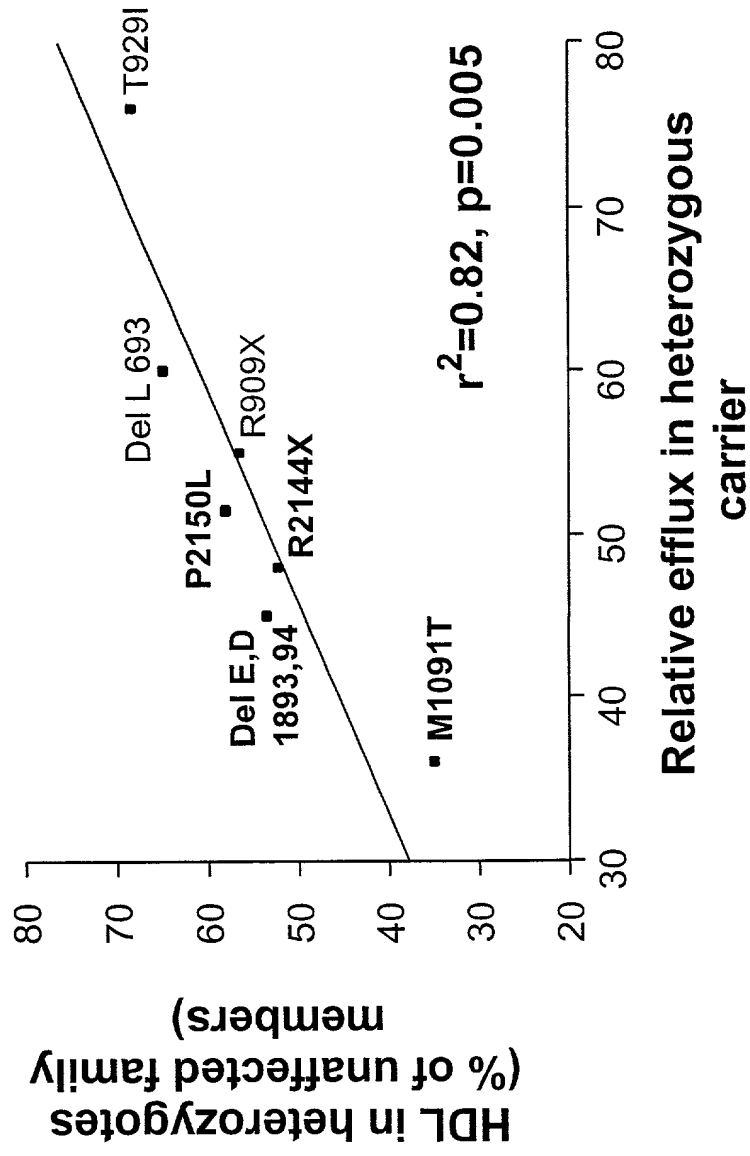


Figure 9

Figure 9

family	mutation	HDL in		HDL in unaffected		HDL in		Age and sex matched		CAD in
		heterozygotes		family members		heterozygotes		population median ^A		
		mean + SD (n)		mean + SD (n)		% of unaffected		mean + SD		
FHA1	Del L 693	0.79±0.20 (8)		1.22±0.35 (11)		64.8		1.39±0.08		-
FHA2	R2144X	0.56±0.23 (12)		1.07±0.22 (20)		52.3		1.34±0.19		+
FHA3	Del E,D 1893,94	0.77±0.24 (8)		1.44±0.38 (9)		53.5		1.30±0.17		+
FHA4	R909X	0.59±0.26 (5)		1.04±0.37 (9)		56.5		1.39±0.24		-
FHA5	M1091T	0.48±0.48 (4)		1.37±0.43 (6)		35.0		1.56±0.05		+
FHA6	P2150L	0.61±0.07 (7)		1.05 (1)		58.1		1.30±0.22		+
TD1	ivs25+1G-->C	0.78±0.06 (4)		1.35±0.29 (70)		57.8		1.22±0.22		-
TD4	del C 6825-->2145X	0.91±0.10 (2)		1.00±0.05 (3)		91.0		1.31±0.16		-
TD5	CTC6952-4TT-->2203X	0.80±0.20 (3)		1.65 (1)		48.5		1.39±0.19		-
TD1	C1477R	0.82±0.18 (9)		1.35±0.29 (70)		60.7		1.37±0.14		+
TD2	Q597R	0.82±0.07 (5)		none available		-		1.39±0.17		-
TD3	T929I	1.01±0.18 (8)		1.48±0.42 (26)		68.2		1.33±0.19		-
TD4	unidentified	0.74±0.05 (2)		1.00±0.05 (3)		73.5		1.49±0.09		+

^A Calculated based on mean the age and sex specific 50th percentile levels in the LRC population

Figure 10

	Missense	Severe	P-value	Unaffected	P-Value	P-value
	Mutations	Mutations	Missense vs.	Controls	Missense vs.	Severe vs.
	(n=33)	(n=42)^A	Severe	(n= 156)	unaffected	unaffected
TC (mmol/L)	4.78±1.30	4.30±0.95	0.08	4.71±1.07	0.76	0.02
TG (mmol/L)	1.77±2.15	1.55±1.01	0.58	1.20±1.03	0.14	0.06
HDL (mmol/L)	0.78±0.26	0.70±0.23	0.18	1.31±0.35	<0.0001	<0.0001
LDL (mmol/L)	3.19±1.10	2.90±0.91	0.23	2.84±0.87	0.10	0.73

^A for TC, TG, LDL measurements, n=41 for severe mutations, 153 for unaffected

Figure 11

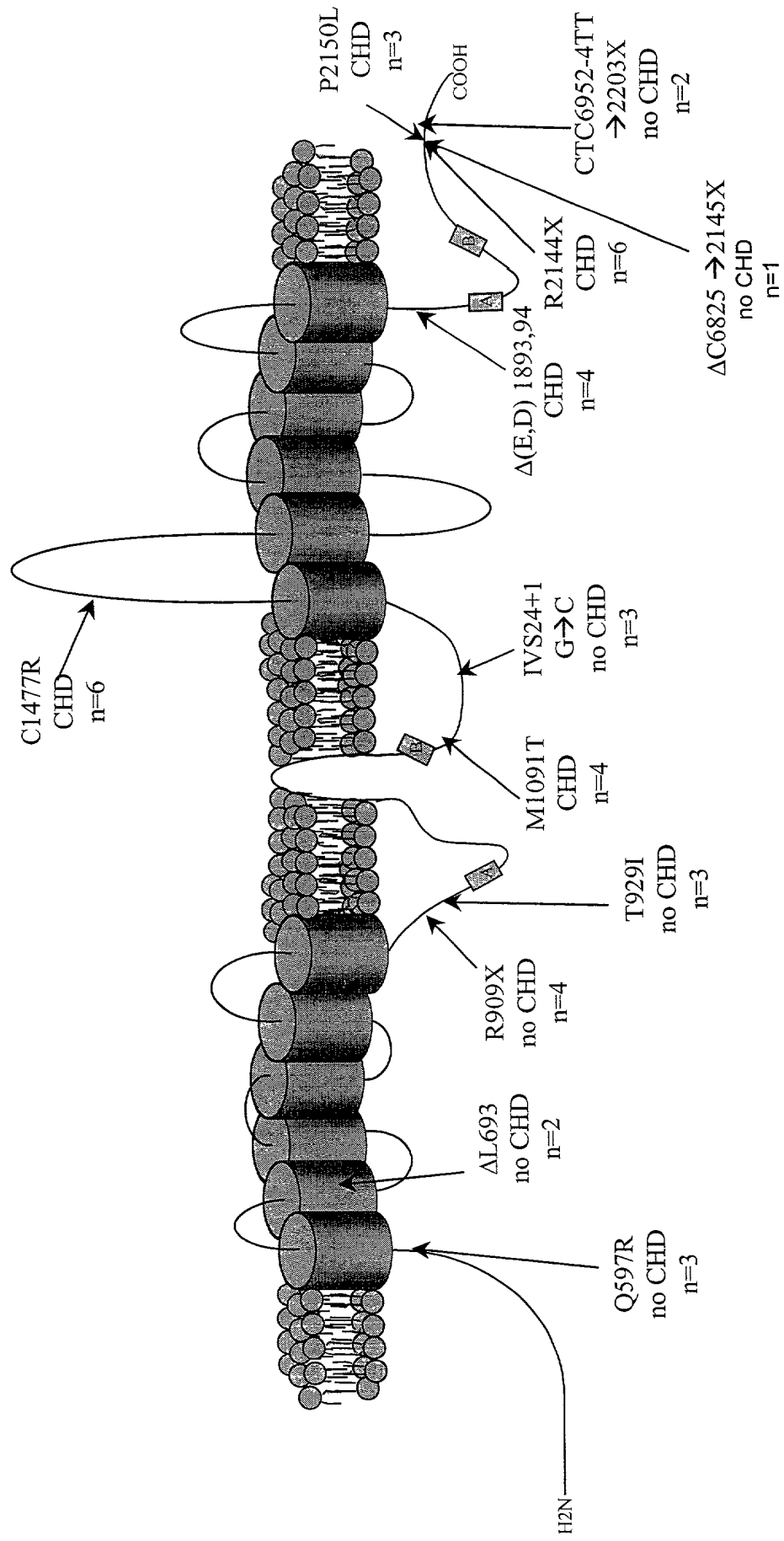


Figure 12A.

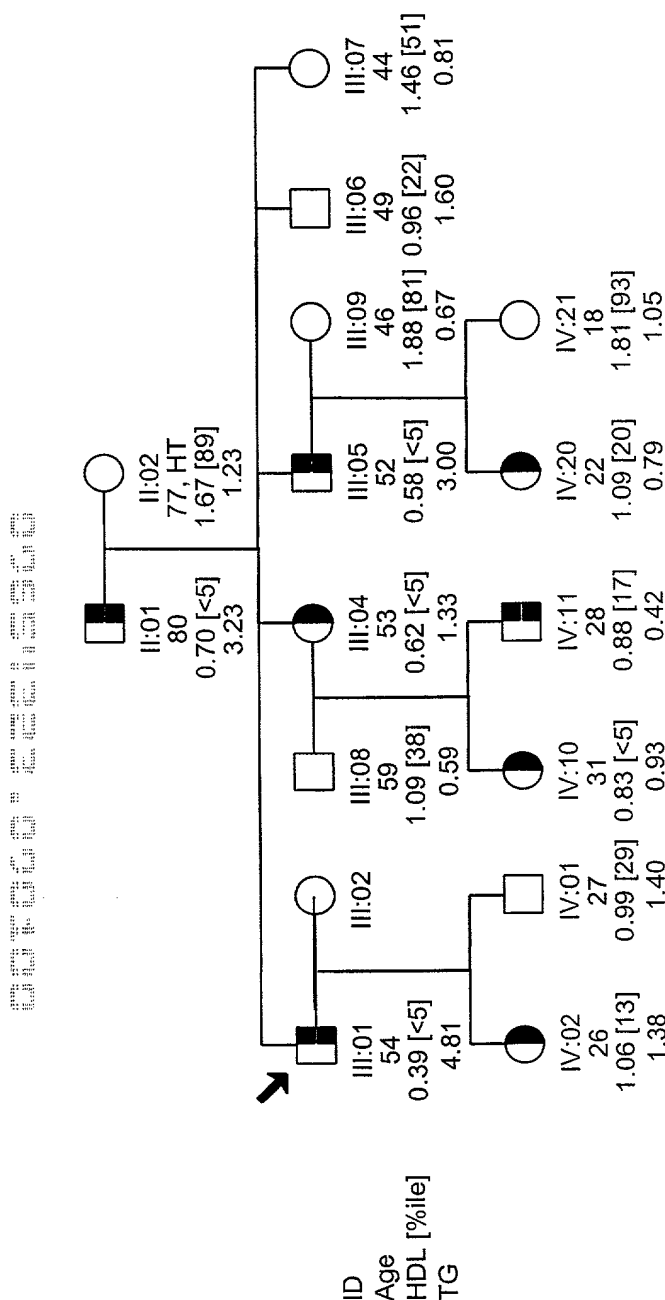


Figure 12B.

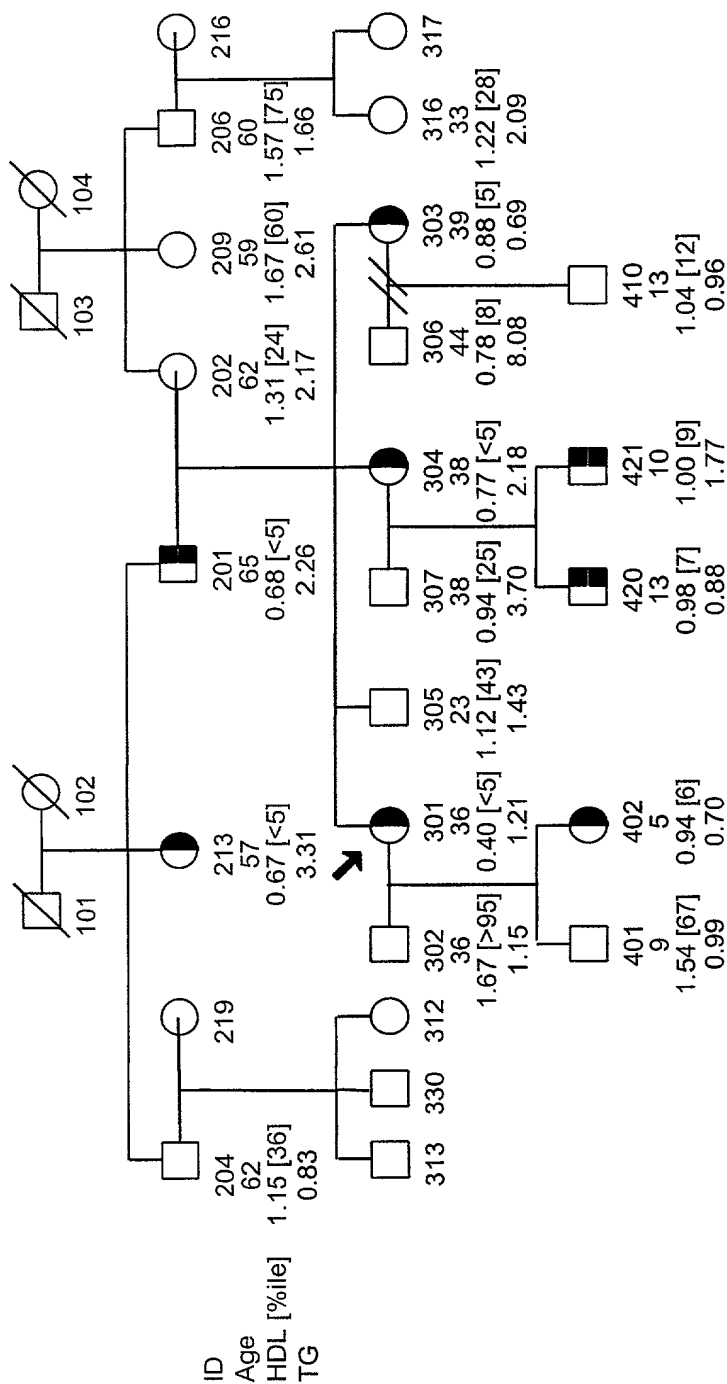


Figure 13

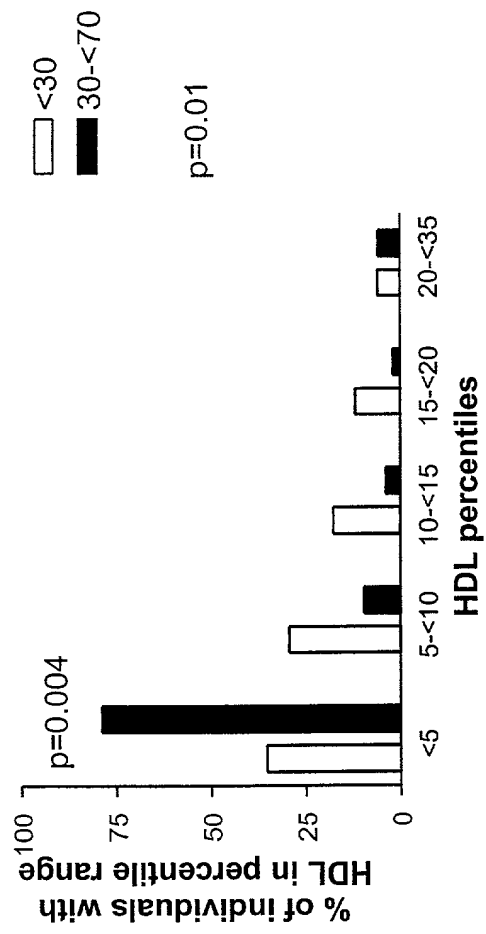


Figure 14

	Heterozygotes	Unaffected	P-value Heterozygotes
	mean±SD (n)	mean±SD (n)	vs. Unaffected
HDL (mmol/L)			
<30	0.91±0.16 (17)	1.26±0.29 (51)	<0.0001
≥30	0.66±0.24 (52)	1.32±0.36 (90)	<0.0001
Change	-0.25	+0.06	0.21
p-value <30 vs. ≥30	0.0002	0.23	
TG (mmol/L)			
<30	1.07±0.96 (16)	0.88±0.45 (51)	0.26
≥30	1.84±1.79 (52)	1.36±1.24 (87)	0.07
Change	+0.77	+0.48	0.97
p-value <30 vs. ≥30	0.03	0.001	

Figure 15A

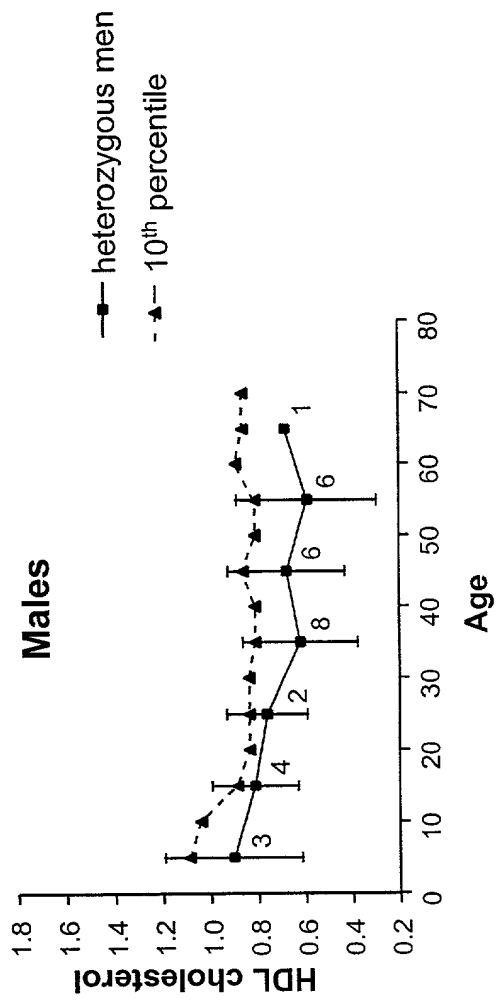


Figure 15B

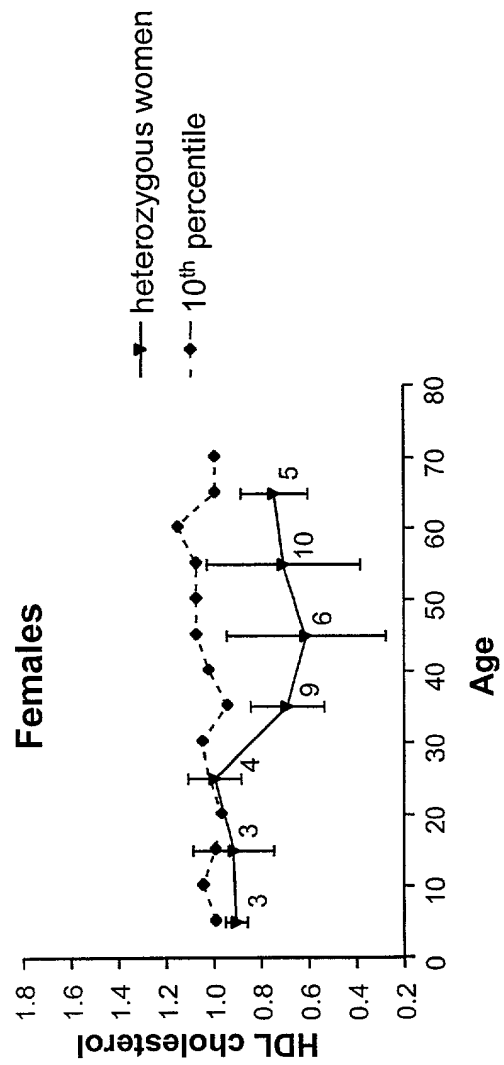


Figure 16A

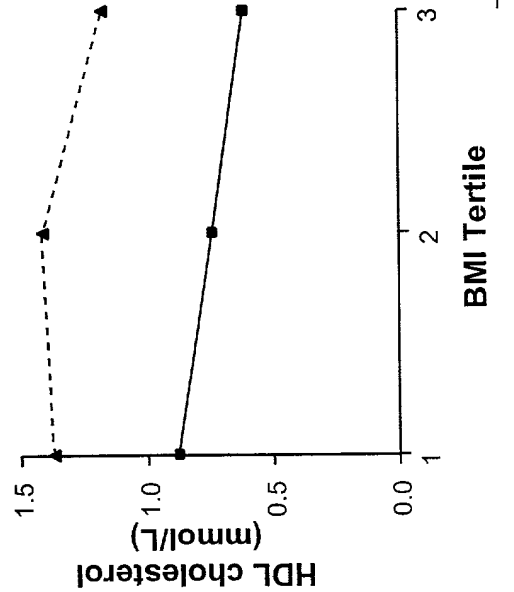


Figure 16B

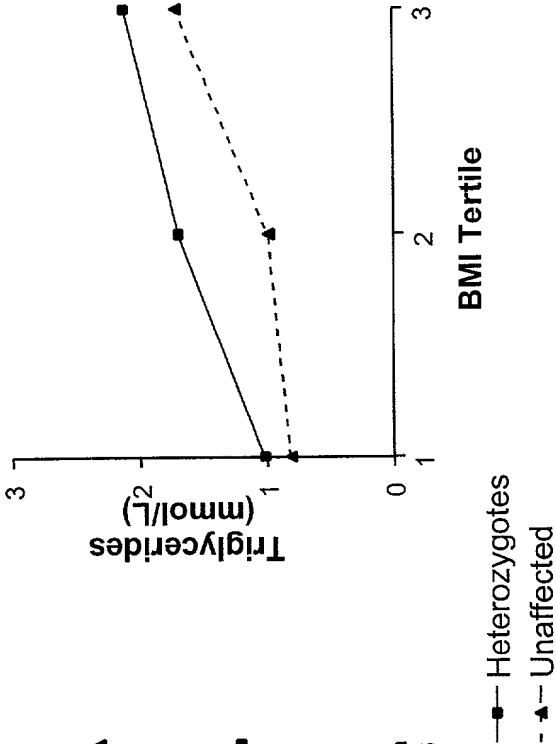


Figure 17

Variant	Amount of each dNTP (uM)	Pmol each oligo	Forward oligo (5'-->3') ^a Reverse oligo (5'-->3') ^a	Annealing temp. (C)	Enzyme	Products (bp): wt "A" allele variant "B" allele	SEQ ID No:
G1051A (R219K)	187.5	20	GTATTTTGGCAAGGCTACCAAGTTACATTTGACAA GATTGGCTTCAGGATGTCATGTTGGAA	60	EcoN I	177 107,70	211 212
T1591C (V399A)	200	27.5	GCTGCTGTGATGGGGTATCT ACCTCACTCACACCTGGGAA	57	Hph I	117, 103, 48, 33 220, 48, 33	213 214
G2706A (V771M)	200	27.5	CAAGTGAGTGCTTGGGATTG TGCTTTTATTCAGGGACTCCA	57	BsaA I	98, 252 350	215 216
A2715C (T774P)	200	27.5	GTGATCCCAGCGTGGTGTGTTGCTT GAAAGGCCAGAGGTACTCACAGCGAAGATCTTGAGGG	55	Hph I	56, 69, 95 56, 161	217 218
G2723C (K776N)	187.5	12	TCGTTTTATTACAGGGACTCCA CAAGTGAGTGCTTGGGATTG	55	Bgl II	269, 80 349	219 220
G2868A (V825I)	200	27.5	CCCATGCACCTGCAGAGATTG GCAAAATCAAAATTTCTCCAGG	57	Bsa I	149, 237 386	221 222
A3044G (I883M)	200	27.5	GAGAAAGAGCCACCCTGGTTCACCAAGAGAGGA AAGGCAGGAGACATCGCTT	55	EcoR V	94, 35 129	223 224
G3911C (E1172C)	200	27.5	GAGCAGTTCGTGATGCTGGCCTGGGCAGCGACCA TCTGCACCTCTCCTCCTCTG	55	BssS I	104, 37 141	225 226
G5155A (R1587K)	200	27.5	CAGCTTGGGAAGATTTATGACAGGACTGGACACGA ATGCCCTGCCAACTTAC	55	BssS I	114, 31 145	227 228
C5587G (S1731C)	187.5	20	GTGCAATTACGTTGTCCCTGCCACACT CCATACAGCAAAAGTAGAAGGGCTAGCACA	60	Mnl I	82, 35 117	229 230
G(-191)C	187.5	24	CAGCGCTTCCCGCGGCTCTTAG CCACTCACTCTCGTCCGCAATTAC	60	HgaI	287, 55, 3 342, 3	233 234

C(-17)G	187.5	18	CTGCTGAGTGACTGAACCTACATAAAACAGAGGCCGGGIA CCACTCACTCTCGTCCGCAATTAC	60	Rsa I	161 124, 37	235 236
C69T	187.5	24	CAGCGCTTCCCGCGGCTCTTAG CCACTCACTCTCGTCCGCAATTAC	60	BsmAI	345 310,35	237 238
C127G	187.5	24	CTGGCTTTCCTGCTGAGTGAC GATCAAAAGTCCCCGAAACC	60	co 0109	284, 175 459	239 240
A(-362)G	187.5	24	ACTCAGTTGTATAACCCCACTGAAAATGAGT TTCTATAGATGTTATCATCTGGG	55	Mbo II	224, 26 134, 90, 26	241 242
A(-461)C	187.5	20	ACTCAGTTGTATAACCCCACTGAAAATGAGT TTCTATAGATGTTATCATCTGGG	55	Hinf I	150, 100 123, 100, 27	243 244
G(-720)A	187.5	20	TCATCTAAGGCACGTTGTGG CCTCAAGCCTGGAGTGACTT	60	Hpa II	450 306, 144	245 246
G(-1027)A	187.5	20	ATGGCAAACAGTCCTCCAAG ACCCTAGCGCTGTGTCTCTG	60	Nco I	170, 41 105, 65, 41	247 248
A(-1095)G	187.5	20	ATGGCAAACAGTCCTCCAAG ACCCTAGCGCTGTGTCTCTG	60	MspA1 I	211 172, 39	249 250
insCCC.T(-1103)	187.5	20	TGTGTGCTCCTCCCTTCCATT CTTGAGGAGCTGTTGGCCAT	60	Mnl I	144 28 11 4 100 48, 28, 11, 4	251 252
insG319	200	27.5	CCCCTCCTGCTTTATCTTTTCAGTTAATGACCAGCCCGG ATCCCCAACTCAAAACCACA	55	Sma I	246 210, 37	253 254
G378C	200	27.5	GCCGCTGCCTTCCAGGGCTCCCGAGCCACACGCTGGG ATCCCCAACTCAAAACCACA	55	Acl I	108, 41, 33, 5 141, 41, 5	255 256

^a Bold indicates mismatch in oligo to create restriction site

Figure 18

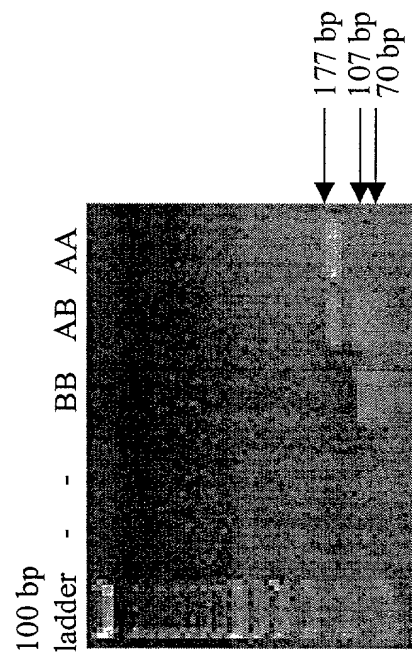


Figure 19

Nucleotide Change	Amino Acid Change	Exon	"B" allele	Frequency carrier allele (%) (B)	N ^a	Frequency carrier allele (%) (B)	Dutch low HDL	Frequency carrier allele (%) (B)	N	Frequency carrier allele (%) (B)	Dutch premature CAD	Frequency carrier allele (%) (B)	N	Frequency carrier allele (%) (B)	Dutch control	Frequency carrier allele (%) (B)	N	Frequency carrier allele (%) (B)	French Canadian
Synonymous																			
G869A	none	6	A	62.5	0.344	32	<i>From Sequencing</i>												
C1331T	none	9	T	26.7	0.133	30													
G1343A	none	9	A	26.7	0.133	30													
T3554G	none	22	G	11.8	0.059	34													
C6842T	none	49	T	6.7	0.033	30													
Non-synonymous																			
<i>REGRESS</i>																			
G1051A	R219K	7	A	46.3	0.254	1588													
T1591C	V399A	11	C	1.6	0.008	1098													
G2706A	V771M	16	A	5.8	0.029	1270													
A2715C	T774P	16	C	0.6	0.003	1250													
G2723C	K776N	16	C	0.5	0.003	1106													
G2868A	V825I	17	A	15.7	0.081	1364													
A3044G	I883M	18	G	23.8	0.136	840													
G3911C	E1172D	24	C	5.3	0.026	1288													
G5155A	R1587K	35	A	44.3	0.259	1566													
C5587G	S1731C	38	G	0	0	558													

^a N refers to the number of alleles screened.

Figure 20

R219K	P-value			P-value		
	AA	AB	BB	AA vs. AB	AA vs. BB	AA vs. AB+BB
n	424	330	36			
MSD	2.70±0.37	2.77±0.37	2.78±0.40	0.01	0.22	0.005
MOD	1.73±0.35	1.81±0.35	1.85±0.35	0.002	0.05	0.001
MI before trial %(n)	48.3 (205)	47.1 (155)	33.3 (12)	0.71	0.12	0.48
events during trial %(n)	17 (71)	13 (41)	11 (4)	0.10	0.49	0.09
total events ^a (%)	65.1 (276)	59.4 (196)	44.4 (16)	0.11	0.01 ^b	0.04 ^c

^a Total events is calculated as the number of events/total number of individuals. Thus, the maximum value for this variable would be 200%, as individuals may have had events both before and during the trial.

^b Odds ratio for BB vs. AA=0.43, 95% confidence interval 0.22-0.85

^c Odds ratio for AB+BB vs. AA=0.74, 95% confidence interval 0.55-0.98

Figure 21

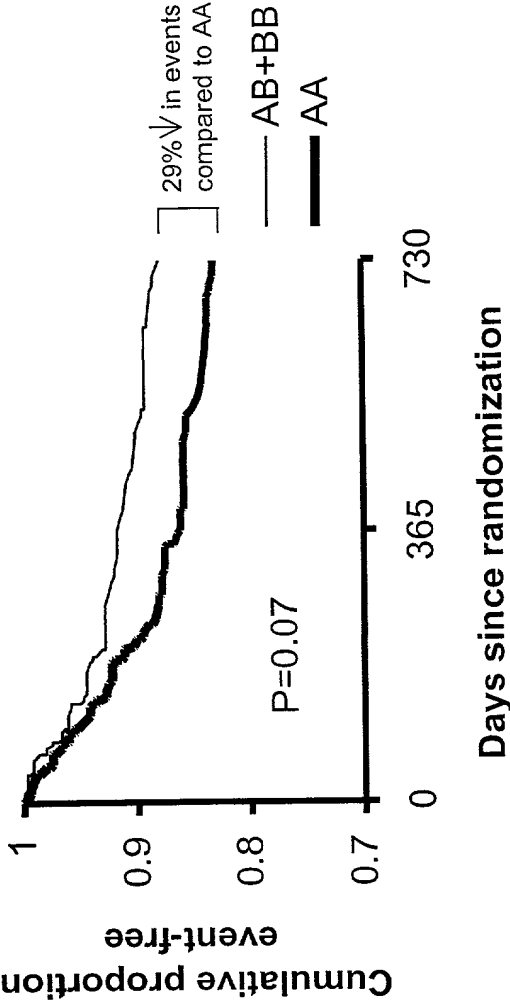


Figure 22

	R219K	AA	AB	BB	P-value		
					AA vs. AB	AA vs. BB	AA vs. AB+BB
n		424	330	36			
Age		57±8	55±8	57±7	0.0007	1	0.03
BMI		25.8±2.6	26.3±2.7	25.5±2.3	0.01	0.50	<i>0.09</i>
Total Cholesterol		6.02±0.86	6.07±0.89	5.89±0.85	0.44	0.38	0.60
HDL Cholesterol		0.92±0.22	0.93±0.23	0.92±0.20	0.54	1	0.81
LDL Cholesterol		4.27±0.75	4.35±0.83	4.33±0.82	<i>0.17</i>	0.65	<i>0.19</i>
Triglycerides		1.84±0.77	1.78±0.78	1.42±0.49	0.29	0.001	<i>0.08</i>

Figure 23

		< median		n	> median		P-value	P-value		P-value		
		mean±SD		n	mean±SD		< vs. >median	AB+BB vs. AA		AB+BB vs. AA		
							< median		> median			
AB+BB												
Total Cholesterol	193	6.22±0.91		172	5.87±0.82		0.0001	0.22		0.43		
HDL cholesterol	192	0.91±0.22		171	0.94±0.23		0.21	0.12		0.37		
LDL Cholesterol	192	4.49±0.84		171	4.19±0.78		0.0005	0.03		0.57		
Triglycerides	193	1.82±0.79		172	1.65±0.72		0.03	0.02		0.85		
MSD	193	2.79±0.37		171	2.75±0.37		0.30	0.18		0.01		
MOD	193	1.83±0.36		171	1.78±0.34		0.18	0.09		0.006		
AA												
Total Cholesterol	207	6.11±0.86		217	5.94±0.84		0.04					
HDL cholesterol	206	0.88±0.20		214	0.96±0.24		0.0002					
LDL Cholesterol	205	4.32±0.77		214	4.23±0.72		0.22					
Triglycerides	206	2.02±0.82		217	1.67±0.67		<0.0001					
MSD	205	2.75±0.36		217	2.65±0.38		0.006					
MOD	205	1.77±0.34		217	1.69±0.35		0.04					

Figure 24

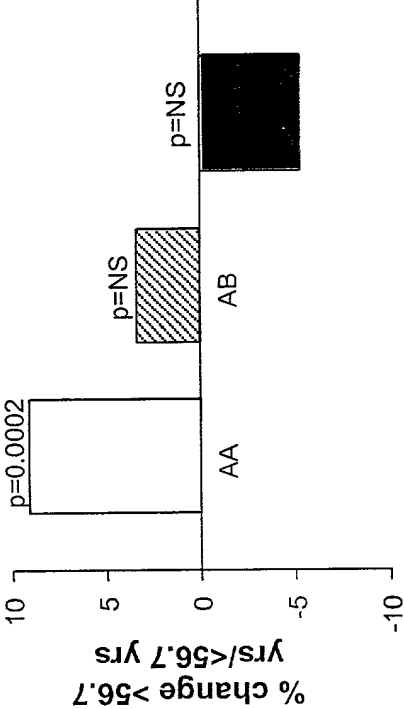


Figure 25A

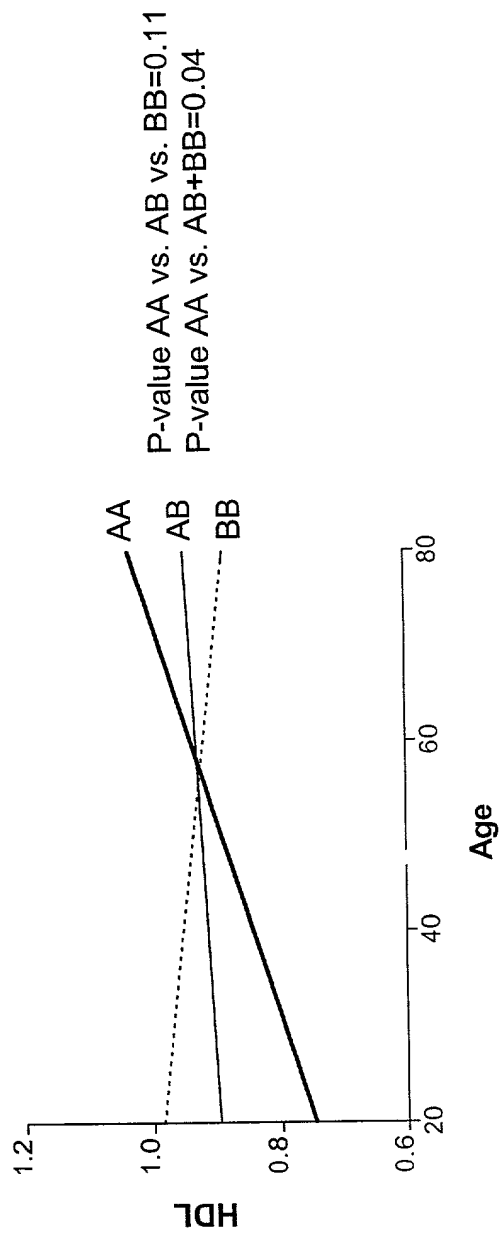


Figure 25B

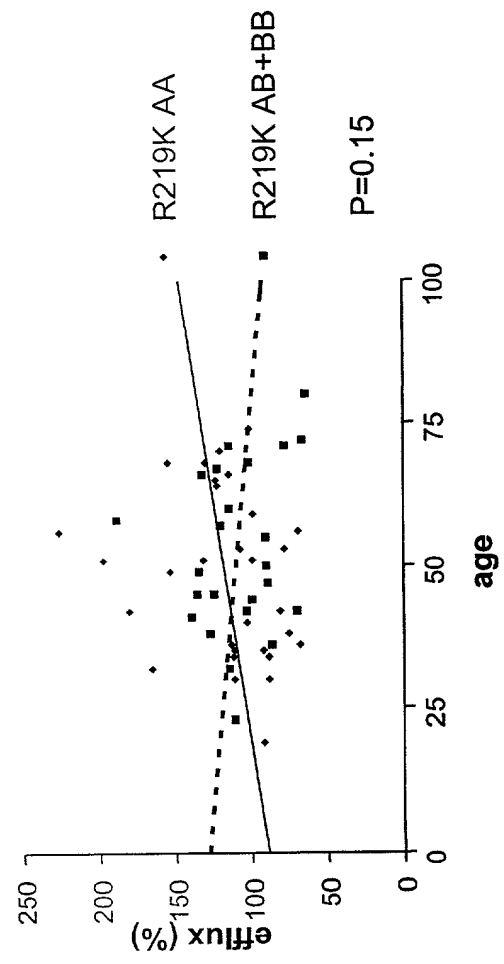


Figure 26A

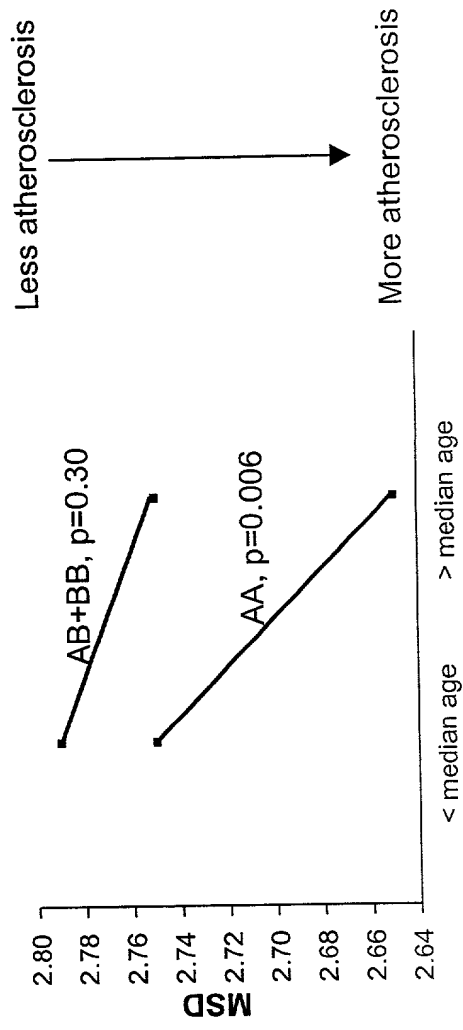


Figure 26A

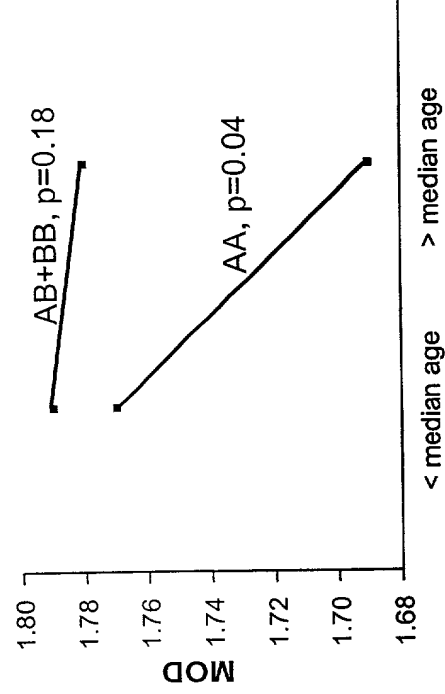


Figure 27

	South African Black ^a	Cantonese	Dutch ^b	P-value South African vs. Dutch	P-value Cantonese vs. Dutch
AA % (n)	1.3 (1)	32.7 (33)	52.5 (180)		
AB % (n)	50.7 (38)	55.4 (56)	45.2 (155)	<0.0001	<0.0001
BB % (n)	48.0 (36)	11.9 (12)	2.3 (8)		
n	75	101	343		
carrier freq.	98.67	67.33	47.52	<0.0001	0.0005
allele freq.	0.733	0.396	0.249	<0.0001	<0.0001

^a Not consistent with Hardy Weinberg equilibrium (p=0.01)

^b Not consistent with Hardy-Weinberg equilibrium (p<0.001)

COMBINED DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled COMPOSITIONS AND METHODS FOR MODULATING HDL CHOLESTEROL AND TRIGLYCERIDE LEVELS, the specification of which

☒ is attached hereto.

☐ was filed on _____ as Application Serial No. _____
and was amended on _____.

☐ was described and claimed in PCT International Application No. _____
filed on _____ and as amended under PCT Article 19 on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information I know to be material to patentability in accordance with Title 37, Code of Federal Regulations, §1.56(a).

FOREIGN PRIORITY RIGHTS: I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

Country	Serial Number	Filing Date	Priority Claimed?
			Yes/No

PROVISIONAL PRIORITY RIGHTS: I hereby claim priority benefits under Title 35, United States Code, §119(e) and §120 of any United States provisional patent application(s) listed below filed by an inventor or inventors on the same subject matter as the present application and having a filing date before that of the application(s) of which priority is claimed:

Serial Number	Filing Date	Status
60/213,958	June 23, 2000	pending
60/124,702	March 15, 1999	abandoned
60/138,048	June 8, 1999	abandoned
60/139,600	June 17, 1999	abandoned

COMBINED DECLARATION AND POWER OF ATTORNEY

Serial Number	Filing Date	Status
60/151,977	September 1, 1999	Pending

NON-PROVISIONAL PRIORITY RIGHTS: I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose all information I know to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56(a) which became available between the filing date of the prior application and the national or PCT international filing date of this application:

Serial Number	Filing Date	Status
09/526,193	March 15, 2000	Pending

I hereby appoint the following attorneys and/or agents to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: Paul T. Clark, Reg. No. 30,162, Karen L. Elbing, Ph.D. Reg. No. 35,238, Kristina Bieker-Brady, Ph.D. Reg. No. 39,109, Susan M. Michaud, Ph.D. Reg. No. 42,885, Mary Rose Scozzafava, Ph.D., Reg. No.36,268, James D. DeCamp, Ph.D., Reg. No. 43,580.

Address all telephone calls to: Paul T. Clark at 617/428-0200.

Address all correspondence to: Paul T. Clark at Clark & Elbing LLP, 176 Federal Street, Boston, MA 02110.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Full Name (First, Middle, Last)	Residence Address (City, State, Country)	Post Office Address (Street, City, State, Country)	Citizenship
Michael R. Hayden	Vancouver, British Columbia CANADA	4484 West 7 th Avenue Vancouver, British Columbia V6R 1W9 CANADA	USA and CANADIAN
Signature:			Date:

COMBINED DECLARATION AND POWER OF ATTORNEY

Full Name (First, Middle, Last)	Residence Address (City, State, Country)	Post Office Address (Street, City, State, Country)	Citizenship
Angela R. Brooks-Wilson	Richmond, British Columbia CANADA	7100 Langton Road Richmond, British Columbia V7C 4B2 CANADA	CANADIAN
Signature:			Date:

Full Name (First, Middle, Last)	Residence Address (City, State, Country)	Post Office Address (Street, City, State, Country)	Citizenship
Simon N. Pimstone	Vancouver, British Columbia CANADA	4746 West 6 th Avenue Vancouver, British Columbia V6T 1C5 CANADA	CANADIAN
Signature:			Date:

Full Name (First, Middle, Last)	Residence Address (City, State, Country)	Post Office Address (Street, City, State, Country)	Citizenship
Susanne M. Clee	Vancouver, British Columbia CANADA	627 E. Osborne Road North Vancouver British Columbia V7N 1M8 CANADA	CANADIAN
Signature:			Date:

SEQUENCE LISTING

<110> Hayden, Michael R.
Pimstone, Simon
Brooks-Wilson, Angela R.
Clee, Susanne M.

<120> Compositions and Methods for Modulating
HDL Cholesterol and Triglyceride Levels

<130> 50110/004002

<150> US 60/124,702

<151> 1999-03-15

<150> US 60/138,048

<151> 1999-06-08

<150> US 60/139,600

<151> 1999-06-17

<150> US 60/151,977

<151> 1999-09-01

<150> US 09/526,193

<151> 2000-03-15

<150> US 60/213,958

<151> 2000-06-23

<160> 256

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 183999

<212> DNA

<213> homo sapien

<220>

<221> variation

<222> (1)...(183999)

<223> N's are A,T,C,G, or other, including no
nucleotide. All K's are G or T. All Y's are C or
T. All M's are A or C. All S's are C or G. All
H's are A, C, or T. All B's are C,G, or T. All
R's are A or G.

<400> 1

gtctataatg gcatgccaca gggctctaaa actttgcagt tttatcatta actcaaatga 60
aatgtatata tgccgctgac tcaacatttt gagagacaac aaatacaatg aatatcaaga 120
tacatatata tataatatgt atttcttttt gagatggagt ttcactgttg ttgtccaggc 180

tctcctagat	tgtcactgaa	taaacattag	cactaactat	ggcaatcaaa	tcacatattg	3480
attgggtgtca	gagaagaatt	gaacattcaa	ctctgaagca	gtgttatttc	ttccatctgc	3540
aaacactctg	cccccatcc	httcctttgtg	tatcctggaa	tccaagtcac	aaataatgat	3600
aggtattttct	gtcaaaggta	tttctcagag	gagtgatatg	taactccctt	tcctctgata	3660
cactgactca	ctaagcagct	acccttgtga	attccaatta	gcaatacttc	ttgctatgtc	3720
tgggtccaaact	ttcagacaaa	cctagtgttc	aggattccta	tagccattta	taggtgtaga	3780
caggaagcat	tcagatatcc	ccagaggtag	ctgataacca	gctgatccat	gactgtctgt	3840
cttgggcttg	gccagcttga	aatcttgaca	ttgtggttct	cncccagaga	aggtgccttt	3900
tggatgtgag	ataaagacat	tatgactaga	tagtgcacac	nnanncacac	acacacacac	4020
taccgaagt	ttgatctgta	aagctgctac	cagcacacac	ttgtcttagg	aaagaatgga	4080
acacacacac	acaattaacc	acagatatcc	tcattgggaaa	atgttcttgg	aagggagttg	4140
agccaagatt	ttatatatag	accacagaag	gtgatgggta	atgttcttgg	aagggagttg	4200
acaggcaata	gctataagtt	aactcaggaa	agcaaaagaa	atccccaaag	agctaaggga	4260
gagggttagag	attctgcttt	ttattagcaa	ttcatagctc	tcaagtttca	tacagtcttt	4320
aaggctcccc	tcttcatata	gaataaatga	aattatttta	taaattgttc	ctcagattcg	4380
tatctgtaca	ttctgggacc	acgagttgta	gcaggatgtg	atcttctcctca	ttctgggcat	4440
ctaagtctta	cagttaagga	cactgaaaca	aacctttagt	cgaataaaga	ttggcacatt	4500
gtttcttctc	ccataaacat	tgaatggtcc	aggaagggcc	aggtgtgggtg	gctcacacct	4560
gtaatcccag	cactttggga	ggccaaggca	ggtagatcac	ccgaggtcag	gagttcaaga	4620
ccagcctggc	caagatgggtg	aaaccctgtc	tctactaaaa	atacaaaaag	gaggagtggg	4680
agtcacagct	tttgcatgaa	tgtccacatt	gggactttat	cttaaatgtg	caagacacaa	4740
ctcttggcca	ggcgtgggtga	accgagattg	tgccactgca	ctccagccag	agaatgcact	4800
ccagcctgag	caacagagcc	agactccttc	tcagaaaaaa	aaaaaaaaaa	aaaaagacaa	4860
ctcttaaaat	ggtgttctac	tgtccttgaa	ccaccatacc	tttgacagaag	tcattctaga	4920
tcataaatct	gaccataaaa	cacagtttgc	caaactggca	actaaactgca	cctattttgtg	4980
attagtttgg	aagaaacttg	aaagtgtccc	ttttaactgt	ccacttattc	cctgcctggg	5040
cacggaattc	tctcccacgc	ttaataacgg	acacttttaa	aaattatttt	tactccccac	5100
agagttggca	gattgctggt	tcagagagtt	aaatggaatg	cctccagtgg	aagtatcccc	5160
ttttctagaa	tggaaagtct	atcttcacag	tgtcataatc	caggtgtccct	gggctgagac	5220
ttcccctgcc	taggcgggtac	cctgggtagc	acagctgaac	tggctgtgaa	ctaaacattc	5280
attttttatt	agcagccagg	ctggacagag	atcacacaag	accaacctga	caacagcagc	5340
atgctgctct	gcttcagaag	taattttttt	tttttttgag	acaagagcct	tgctctatca	5400
cccaggcaaa	gtgcagtggc	atgatctcgg	ctcactgcaa	cctccacctt	ctgggatcaa	5460
gcgattctcc	tgcctcagcc	tcccagtag	ctgggattac	aggcgcacgc	caccatgcct	5520
ggctaatttt	ttgkaacttt	gtattttttt	tttttttgag	atggagtcct	cgctctgtcg	5580
cccgggctgg	agtggcctga	gccaccgcac	ctggcccttc	agaagtaatt	tttaagggaag	5640
gatcttgtcc	tgggtggggg	gttgggcaag	gctgtgaaaa	gcaaacaaaa	atcacatgtg	5700
gcctaataag	aggccagtgg	aaatacacat	gatgaaaaag	aaacttacaa	aagcacatta	5760
ttaattttctg	aacatgctaa	taccatccaa	taacaataag	atctaataat	tattgaatgc	5820
tgattactca	tcagctacat	tccaagtact	tctacatgta	cagttatgtg	acacataaga	5880
tgtttaggtc	accactgaca	gaatgtatga	tggtagcccc	ataagattac	aaaaattcct	5940
attgcctagt	gacatatgta	gcaagatgca	ttactcacgt	gtaataagta	atgtgatgct	6000
gctgtaaaaa	aacctactgt	gctgccatgc	ctatagaagt	ctaggacggg	ctgggcacag	6060
tggctcatgc	ctgtaatccc	agtacgggtg	gaggccaagg	caggccgatt	gcttgagccc	6120
agaagttaa	gaccagccgg	ggcaacatag	tgagagcctg	tctctacaaa	aaatacaaga	6180
acaaattagc	caggcatggt	ggcgcatgcc	tttgggtcca	gctactcagg	aggctgaggc	6240
aggaggatca	cttaagccca	ggaggttgag	tctgcagtga	gccatgatgg	tgccactgca	6300
ctccaaactg	gagagagagt	gagaccgggt	tttaaaaaaa	aaaaaaagcc	taacaaataa	6360
aagtatgtat	agtacagaat	acttggtaat	gataaacagt	atgttactgg	tttatgtgtt	6420
tactatactt	tttatttttt	agagtatact	cctacttatt	taaaaagaaa	aaagaaagtt	6480
aactgtaaaa	caggcttaga	caggctccttc	aggaggtatt	ccagaagaaa	ggcattgtta	6540
tcacaggaga	tgaaagctcc	atgcatgtta	ctgcccttga	agaccttcca	gtgggacaag	6600
atgtgaggtg	gaagtctatg	atagtgatga	tcctgaccct	gtgtaggcct	aggctaatat	6660
gtgtgttttt	aacatattag	tttttaacaa	caaagttaa	tcagttaaaa	caattttttt	

aatagaaaa	agctcataga	ataaggatat	naaagaaang	acaatgtttt	gtttgtttgt	6720
ttgtttgttt	tttgagacag	agtctcgctn	ccatcacctg	ggctggagt	cagtggcaca	6780
atcttggtct	agtgaacct	ctgattccca	ggttcaagta	attctcatgc	ctcaacctcc	6840
caagtagctg	ggattacagg	cgcccatcac	cacaccaaac	taacttttgt	cttttttagta	6900
gagatggggg	ttcatcatat	tggccaggct	tgtcttgaac	tcctgacctc	aggtgatcca	6960
ccctcctcgg	cctcccaaag	tgctgggatt	acggttgtaa	gccatcgtgc	ccggccagaa	7020
agaaaaatgtt	tgtacagtat	gtacgatgta	tttgttttaa	gctaagtgtt	attatgaatg	7080
aatcatgtgt	ggtgtagtgg	ctcatgccc	taatcgcatc	actttgggag	accgaggacg	7140
gaggatagct	tgagtccaga	gttctagacc	agcctgaaca	acgtgggtgag	acctcgtctc	7200
tacgaaaatc	atcaaaaatt	atccatgcgt	gggggtacat	acctatagtc	ctagctactc	7260
aggagactga	ggtgagggga	ttgcttgaac	tcaagaagtt	gaagctgcag	tgagctataa	7320
ttgtaccact	gcatttcggc	ctaggtgacc	ccgcctcaaa	gaataataat	aaaagctggg	7380
cacggtggct	cacgcctgta	atcccagcac	tttgggaggc	caaggcgggc	ggatcacttg	7440
aggtcaggag	tttgagacca	gcctggctga	catggtgaga	ccccgtcgct	actaaatata	7500
aaaaatttagc	cggtcgtggt	ggtgtgtgcc	tgtaatccca	gctactcagg	aagctgaggt	7560
gggagaattg	cttgaacctg	ggaggcagag	ggtgcagtga	gccgagatag	cgccactgca	7620
ctccagcctg	agcaaaaaac	aataataaat	aaataagttt	gaaaatttaa	aatgtttata	7680
aaattaaaaa	gttacagtga	gctaagattt	agtattaaag	aattttttta	taaacatgta	7740
gtgtaactct	acagtgttaa	taaactctac	agtagtgtac	agtgacatcc	taggccttca	7800
cattcactca	tcactcactg	actcacccag	agcaacttcc	agtcctgaaa	gctccattca	7860
tggtaaagttt	cctatacagg	cggtgccattt	aaaaaatctt	ttagggccagg	cgcgccggct	7920
cacgnctgta	atcccagcac	tttgggaggc	tgaggcaggc	ggatcacgag	ctcaggagat	7980
cgagaccatc	ctggctaaca	cggtgaaacc	ccgtctctac	taaaaataca	aaagattagc	8040
tgggtgtggt	ggtgggcgcc	tgtagtccca	gctactcggg	aggctgaggc	aggagaatgg	8100
catgaaccca	ggaggcggag	cttgacagt	agctgagatc	acaccactgc	gctccagcct	8160
gggccacaga	cgaagactcc	atctcaaaaa	aaaaaaaaat	cttttatact	atattcttaa	8220
tgtacctttt	ctatgtttag	atacacaaat	accactgtat	tacaattgcc	tactgtattc	8280
agtacagtaa	catgctatat	ggttacgtag	cctaggaaca	ataggctata	ttgttcctag	8340
gtatagggat	gtggtatcct	ataccatcta	ggtttgtgta	agtttatttt	atgatgtttg	8400
catgacgaca	gagtcaccta	aggactcatt	tcttagaata	tattttagtg	taagcaatgc	8460
atgactctat	tgactcatga	attcttacca	cagacctatg	gggcagtact	attgttacct	8520
tcattttata	aatgataaaa	ctgaggtaca	gagacagtaa	ataacttgac	cacgggtcatt	8580
cagctactca	agagtcaagg	ctgggattta	aaaccagatc	acatggtttc	agagtgttca	8640
cacttaccta	ctatactgtc	tcaagagcaa	ggatgttttg	gttcacttga	caaatgaaga	8700
tagggacctc	tttcattata	agcctatttt	aggctaaaat	agaaggggag	gggacacagt	8760
gaatccaggc	cttctggcat	ggctcctcag	ccctttctga	gctggcctgg	gacagccttc	8820
ctacctcact	gatgccactt	cctactgagc	gactttcctg	cctacctcac	tgatgccact	8880
tcctactgag	cgactttcct	gggctccaga	cccagtaagc	gactttgcct	gcacccacct	8940
tattttgctc	tactccctgt	gcttttatgc	ctttacccat	ctgccctgga	aagctcttct	9000
aacctttgaa	tgggtaaagg	cataaatgta	tgctggagaa	atcctcagct	cagggccagg	9060
cacgctggct	cacgcctgta	atcccagcac	tttgggaggc	cgaggtgggt	agatcacctg	9120
aggtcgggag	ttggagacca	gcctgaccaa	catggagaaa	cccctgtctc	tactaaaaat	9180
acaaaattag	ctgggcgtga	tggcacatgc	ctgtaatccc	acagctactc	gggaggctga	9240
ggcaggagaa	ttgcttgaac	ctgggaggcg	gaggttgtag	tgagccaaga	ttgcgcctct	9300
gcactccagc	ctgggcgaca	gagctagact	ctgcctcaaa	aaaaaaaaac	aaaaaaaaaga	9360
aaaaaaaagaa	aaaaaagaaa	tcctcagctg	agttgtcaac	tcctctttga	aactttctca	9420
gacctttcag	gctgagtcgt	cgtcattttg	tgtcttctca	gttctctggc	tctaccttct	9480
tcatagcttg	tttcatgtaa	tgttaattct	tacttgcttt	ctccctcttc	taagctgaga	9540
gctacttcaa	agcatgggta	ggacctagca	cggtgtatgg	gacatgggtg	gtaccccgta	9600
aatgtttact	gaaaaaaaaa	tgcctaaagc	aattgttaac	atcatcagat	agataattat	9660
gggcattcag	agattctgtc	ttcaagctta	tataaagaac	ttatttttgg	ctctaattat	9720
cctgataatt	ttctcattac	tttcaattat	tgtggcttgt	ggatcaattg	ttgacatttt	9780
ataaacattt	cactatttga	caatgatgat	actaaaatac	gaattaagca	accattctaa	9840
agatagtgat	gatgataaca	tatacgctgg	taacatcttt	attttcagcc	gtatcatgga	9900

atcctctgtt	tccattctgc	taggtaggca	ggtatgcagg	tagaacttgt	gagaggatat	9960
gatttttgtt	tccatcttag	atatgacagg	aacttggaa	ttttgacata	aatgacgaac	10020
atccgggatt	cttaacaat	ctttaaaaat	ggaatgcctt	aaaagctggg	cgagtggt	10080
cacgcttata	atcccagcac	tttgggaggc	tgaggcaaat	ggatcacttg	agttcaggag	10140
ttcaagacca	gcctggccaa	catggtgaaa	ccccatctct	actacaaata	caaataattag	10200
ccgggcgtag	tggcaggcgc	ctgtaatccc	agctacttgg	gagcctgagg	cagggggaatt	10260
gcttgaaccc	aggaggcctt	ggagattgca	gtgagctgag	actgcgccat	tgcactccag	10320
cctgggcaac	aagagtga	ctccatctcc	ggaaaaaaa	aaaaaaaaa	aaggaatgcc	10380
tttgggaata	atttatatt	aatttatgta	taacatatag	acaaaccatt	agtttgtctt	10440
atattttact	aaatataaat	ttagtaaata	taaatatatta	ctaaatataa	aaactcttag	10500
attttactaa	agagttacaa	ctaattggcc	tggcgtgggtg	gctcacacct	gtaatcccag	10560
cactttagga	ggcagagggtg	ggccgatcac	gagggtcagga	gatcaagacc	atcctggcta	10620
acacggtgaa	actctgtctc	tactaaaaaa	aaaaaaatac	aaaaaattag	ccgggcggtg	10680
tggcaggccc	ctgtagtccc	agctactcag	gaggctgagg	caggagaatg	gcgtgaacta	10740
agcagaggct	tccctaaaag	tgatcttcag	gataaaggca	gaggaagagg	ctccatgact	10800
gggattgggtg	ttgaggagag	ccagagaagc	aagctacaga	aaagagaaaa	aattaatatg	10860
caagagagta	aacaacacga	aggaaaagaa	cccagtgtgg	aaacactaca	cgtgagaaaag	10920
gtgtctgtaa	ggatgttcta	caaagcaaat	gcttggatat	taattcattg	cagcaggaga	10980
tggtaagcct	catgataaag	aaggagaaaa	aatcaagtca	agggctctga	ggtactgacc	11040
caggataact	tgactatgcc	agcaactgtt	tagggggaga	tttgagctac	acttgtagca	11100
aaggcaaat	ctgtaattag	ttgtaactct	tttttttttt	gagatgggtg	ctcgtctgt	11160
ccccaggct	ggagtgcagt	ggtgtgatct	tggctcactg	caagctccgc	ctcctgggtt	11220
caagtgattc	tccagcctca	gcctcccaag	taattgggac	tacaggcatg	caccaccatg	11280
cccagcta	ttttgtactt	ttattacaga	ccatgttttg	ccatgttcac	caagctggtc	11340
tcaagctcct	gacctcaagt	gatccgtccg	cctcgccnt	cccaaagcgc	tgagattata	11400
ggcctgaacc	accgcgcctg	gcctaaagag	atctaattct	tagcaaagtt	tcaccaggga	11460
gtctctcctc	acccccaccc	catccttccc	acaaagaatt	agaacaatgt	ccctactacc	11520
cctgctgtat	ctctgacttt	ttacttttaa	tctcagcaga	atattttact	aaatgttttg	11580
atgtggttat	ataaaatcat	ccctgctgac	aaggaaacac	tttttgaaaa	aagttttcat	11640
tatcaaacag	taagtacagc	tgactgccgt	gacctttaac	ccatttctga	gtctccccctc	11700
attggacttg	ggtggagggg	actggtacca	ataaagtcaa	atgcttaata	atttatgcaa	11760
gtgcttgaag	aaatttgaag	ttgaatat	ctatcatctt	gaaatggaga	aagaatctgt	11820
aaacagcaaa	gccagacgcc	ctaaaggaaa	agatttacag	attaaaaata	gattgcaatc	11880
tggtaaaaat	atttgcaaca	catgtaacag	tcagaaagtt	gaaacacttg	gtttaacawg	11940
agcttttawc	agataaataa	ggaaagaata	aacattggat	tttaaaccacc	gataaacatg	12000
aaaagatgtt	taatctttat	ttttatttaa	tccatattat	ttttcagttt	aatcaaagaa	12060
aatataaatc	aaaacaataa	tacattttat	atataatat	atataatat	atataatat	12120
atataatatag	taggcataag	ttaaagactg	ataagactgt	tgaaaaggga	tgaaaaacta	12180
ggcttactca	taccaatata	tatctattag	gatggctaaa	gtaaaaaata	ctgaaaaatat	12240
caagtgtctca	aaaggatatg	gagcaattgg	aaccctcaga	catcgctgrt	ngagaaaaca	12300
aaatggtaca	gccaccctgg	agaacagttt	agctgtttct	tgtaaagtta	aacatgcgct	12360
taccatatga	ctcagcaatc	tcactcctgg	gtatttatgc	taggaaaagg	aaaatttata	12420
cttgacacac	aaaaacttgt	aagtgaatct	ttatagcagc	tctattcata	actgccaaaa	12480
actgagagaa	aatgtccttt	aatgtgtgaa	tggataaacc	aactgtgcaa	catccatgta	12540
atgaaatact	acttagcaat	aataataata	atattaaaaa	ccccagaacc	attgatgcat	12600
gcaacaaata	tggataaatc	tcaaaagcat	tatgctgagt	aaaagaagtc	agtctgaagg	12660
atttcatact	ctaggattcc	atttatataa	cattattgaa	acgacaaaat	tatggggaca	12720
gagaatagat	cagcggttgc	caggggttta	ggtgtgtgga	gagggtgtgg	ctataaagaa	12780
catgcaaggg	aatttttttg	ggagatgaaa	tggatctgta	tcctaattat	ggtcatggta	12840
acacaaatct	atacatgtgt	ttagattcat	agaactgtat	accaaagaa	aaaaagtc	12900
ttttactctc	ttaaaatgaa	aaaagaaaaa	gcctgggcat	tctaaccact	tgtttgtgag	12960
agtacacatt	gataccaagt	tttatgggtg	gcaattttgc	tataaatacg	gaaagtttgt	13020
ctgttctatt	attcagcaat	cccagttttg	caaaactatg	ctaaagaatc	tttgggggccc	13080
gggcacggtg	gttcacgcct	gtaatctcag	cactttggga	ggccgaggtg	gacgatcacc	13140

tgaggtcagg	agtttgagac	cagcctggcc	aacattgggc	aaccctgtct	ctactgaaaa	13200
tacaaaaatt	agccgggcat	ggtggcgcat	gcctgtagtc	cgagctactc	gggaggctga	13260
ggcaggagaa	tcacttgaac	ctgggaggca	gaggttgtag	tgaactgaga	tcgtgccacc	13320
gcactccagc	ctgggcaaca	gagtgagaat	ccgtctcaaa	aaaaaaaaaa	aaaaaaaaaa	13380
caaaacaaaa	caaaaaactt	tgtgtacgtg	tgcaaagaga	atacaaagat	gatcatggct	13440
gcatttttta	aatgactata	aaaaagaggt	acaaccagcc	aggtaaagtg	gtgtgcacct	13500
gtagtcccag	ctactcggga	gggtgaggtg	agaggaacac	ttgagtccag	gagtttcagg	13560
ccagcctggg	caacatagtg	agaccctgt	cccaaaaaca	aaacaaaaca	ccaaaatgtc	13620
tatctgtagg	aatttgtttt	caagttgtga	tacataggta	cagtgaata	ttatacattc	13680
atttaaaaatg	atgataaaat	ctgtatttgt	ttacatgaaa	aactgtccac	tataatgtag	13740
tgaataaat	agattacaaa	caatatatat	ggaataaact	tgtttagaaa	caatttctag	13800
aagaaggtag	aatggacaga	attatctctg	ggaagtgggt	ttataatgat	tctcattttc	13860
ttctttgtat	ctttttcata	gtctttctac	ttttgttatg	tctggacatt	tgattatgag	13920
catgtattac	tgatctattt	taaaaattga	ttttaatttt	tacaaaaact	catgtaaagc	13980
ttgaagggtc	gcatttttaga	ccatgttaaa	attttccctgg	atcaaaacag	acttattcaa	14040
atatcttgta	ccctgtcctc	caaaattgcc	tgccaaaata	cactacaaaa	gagagcattt	14100
agctgcatat	tttttggaact	gctgagatca	acaatattat	ttaccatggc	ttaaattttt	14160
accctccagt	atgtgtgggt	tacaaacact	cttccacatt	tttgaggcat	tgcttttgat	14220
attttaaatg	taaattcagc	tgtgcgcggg	ggctaacgcc	tataatccca	gcactttggg	14280
acgctgagga	aaggatcact	tgaggtcagg	agtttgagac	cagtttagct	aacgtgggtga	14340
aaccccgctc	cttttaaaac	tacaaaaatt	aaccgggcat	ggtggcaggc	acctgtaatc	14400
ccagctactc	aggaggctga	ggcaggagaa	tcacttgaac	ttgggagaca	gaggttgcag	14460
tgagccgaga	tcatgccact	gccctccagc	ctggccacag	agcgacactc	catctcaaaa	14520
aaaaaaaaaa	aaaaaaaaaa	aaaggccagg	cgcagtgggt	cacgcctgta	atcccagcac	14580
tttkggagcg	ccaagggtgs	sggatcaact	gaggttgagg	gttcacgacc	agcctgacca	14640
acatgcagaa	accygtctc	tactaaaaat	acaaaattag	ccgggtgtgg	tggtacatgc	14700
ctgtaatccc	agctactcgg	gaggctgagg	caggagaatt	gcttgaaccc	aggaggtgaa	14760
ggttgtgttg	accagagatc	ccgccattgc	actccagcct	gggcaacgag	caaaactctg	14820
tctcaaaaaa	accgaaaaaa	ttcccccaaa	aaaccaaaaa	aaaacagcaa	caacaaaaaa	14880
atcaataaat	gtacctgtgt	tagcataaag	cataattata	tgcatatggg	gattgggagg	14940
atgaaatgga	aaggttattt	attactgact	tcagaaatta	tgtcctgata	gattgatggg	15000
tgatttaaat	ataacttctt	gtcaagcatc	tgtttttaga	atcaaattac	tatgactctg	15060
cagtttcctt	gaaatctcat	agtatcacat	ctctgtttgc	ctttgcatgg	ttttaagaaa	15120
atgaggagtg	tgaaacttca	aacttcgttt	tcattgtatt	acatttttga	atgacacact	15180
ggtcattttc	tagaaatata	agggtgacaaa	tatttcacag	aaacataagg	tgctattatc	15240
tcattcaatc	ttagggtcact	caaaactctt	tctctccac	acattgaaga	ttcatttggg	15300
aatgcttttg	tcttattgtg	cacccccagt	gaaggggtgg	taagtgtttt	tcattttgct	15360
tcctttgttt	atctacaggg	ttccattcaa	taaacaaagg	gacttgggtc	aaacttcagg	15420
ctcttatggg	tttgatgta	atcttttggtc	tcatttttag	ttaccaacag	agagtgttgc	15480
ttctgacctc	tttgactctt	ccctgctgaa	tttactatgc	ctttgatact	tgtgaagggg	15540
gagattttcg	aggagtactg	ttgtttttgt	tagaggttgt	aatgtctttc	ttcgctttgt	15600
gattcaagtt	gtgttcagtt	acaatcataa	gcatgtgcct	aaaaaaatca	gatgcaaaact	15660
agcaaaagta	gaaactcagg	gtgacagtct	ttaagaaaag	atgcaattct	tggggctggg	15720
tgcgggtggc	tatgcctata	atcccagcat	gttgggaggc	caagggtggc	agatcgcccg	15780
aggtcaggag	ttcgagacca	gcctggccaa	tgtggtgaaa	ccctgtctct	actaaaaata	15840
caaaaattag	ctgggcgtgc	tggtgggtgc	ctgtaatccc	agctactcag	gaggctgagg	15900
caggagaatt	gctggaacct	gggagggtga	ggttgcagtg	agcctagatt	gcgccattgc	15960
actccagcct	gggcaacaag	agcgaaactc	tgtctcaaaa	aaaaaaaaaa	aaaaaaagat	16020
gcaactctta	ttactgacac	agaaatgaaa	atttagttac	atagtattgt	aaaaggacta	16080
tcagctaggt	ttagccttac	caagatttag	gtaattcatt	tcctgctaca	ctcatattct	16140
cagccacttc	cctcatcaca	ttttcagggt	gcagtatata	atagcgtcaa	ctcgtgtaat	16200
tccccctact	ccccatgaac	ttctaggcca	agggggccaca	cgggggtggg	catatagtat	16260
aaaggagtaa	ggcagactgt	tggagaaaaa	cagggattag	gccaggtgcg	gtggctcacg	16320
tgtgtaatcc	cagcacttta	ggaggctgag	gcggggcgat	cacgagggtca	agagtttgag	16380

tggtcaggct	ggtctcaaac	tcctgaccte	aggtgatctg	catgcctcgg	cctcccaaag	19680
tgctgggatt	acaggcatga	gccactggtc	ctggaccaa	gcctcttgat	acatttttgct	19740
aaactatcag	aatgttgggc	ctctttgcac	cctactaaca	ccctcagttt	ttatgagtat	19800
gattacatta	ttttccctct	tttccccaac	ctcctattct	attttcctta	gataacctta	19860
gactgcaccc	ccaacttagc	tttctgttct	cattcaggca	gtgtgtccag	ggcttccaat	19920
cccagtgggtg	agtaggcaga	tgtgggggct	tcctgttgaa	ctctgttagg	taattgagcc	19980
actcaagggg	aaaccagctt	tgggggtccgc	tctgtggcaa	gccaggact	ctatngaag	20040
gagctcattc	ctcactaatt	tacacaccag	gacaagtagc	cangctaaga	aattgctcta	20100
atttcactgg	tcgtcagtc	acaaaagaat	gaagatacag	aaggaaaata	aaaggcctgc	20160
tcccacagtg	gggtcatagc	acatctccac	atttaacctt	tgacaaggca	aaagtatttc	20220
ttctgtccag	acaattccag	aagtcataag	gaatactggg	caaggatggg	tcaaaacaac	20280
acagactatt	gagggcaaca	aaggaaaaac	catgcactcc	tgaagaaata	gatcccaaaa	20340
tgctgaacga	gaaagaggat	tgtctctcta	gaagtcta	agaaaacagt	atctggaaca	20400
ttaaacgtag	gatttttttt	ttttctaata	acttactctc	ttattaatat	gtcagagaaa	20460
agaatagctc	ctggctaaga	aatacaacag	tccctcatcc	cagaaatcac	agccaggaat	20520
atggattctt	aagtgtcaaa	aaagtactct	gaataaggaa	agaaacgcag	atgcactact	20580
tctattataa	taagcgcttg	atttctttta	atcccttgag	tccaattatt	ttcttggcaa	20640
atttaaggg	actgactgct	tctgtgaatc	tattgttaca	cttgataatg	gatctgagtt	20700
gggggtaata	tttgctatc	aatttggata	cttaaaaatc	tctctctctt	tcttcatata	20760
ccctctatct	cacaactttc	catttaata	gggaagtga	tttctttttt	ctgccccctc	20820
ttcctactgc	ttctagaata	aaagcataac	aggacgacag	gagtggagat	gagaggaagc	20880
atttccaagc	aatgggaaag	tatgatgaga	gtcgtgagtt	ggtagaatgg	gggtgagtaa	20940
gggggtgggg	gtggatggat	ggaagaggat	ggaggaggaa	gcaggtcata	tgatcaggct	21000
tcaaaggcct	tccaatatct	tgttttcaga	ggcttctgag	gtcttcttac	agaactttga	21060
gggtattttt	cctttgtgag	caaaggagaa	ccatagtaaa	ttttgaggg	tgggctgggc	21120
gcgggtggctc	acgcctgtaa	tcccagcact	gtggcaggca	gaggcagggtg	gatcatttga	21180
ggtcaggagt	tcgagaccag	cctggccaaa	atgggtgaaac	cccgtctcta	ctgaaaaaac	21240
aaaaagttag	ttggatgtgg	ttgcatgtgc	ctataactcc	agctactcgg	gcagctgagg	21300
cagcagactt	gctggaacc	gggaggcaga	tgttgcaatg	agttgagatt	gtgccactgc	21360
actccagtct	gggagacaga	gtgagactcc	atctcaaaaa	aaaagaaatt	tttttttttg	21420
gagggagaaa	atatatgatt	agattgtttt	tgtgttggtt	tgcttgtttg	ttttttttcc	21480
caagcaaaaa	atcactttac	tgcaataggg	aagacaaata	ggaaggggaa	gaaactagag	21540
acaggaacag	cagtttgag	gttctgcaat	acagaagccc	agatgtttgg	cttggaactta	21600
gacactggga	atgaaaaata	aatgatgaat	taaaaaaata	aaatatttgg	gaggtatact	21660
tgacctgacc	ttgggtgctt	ttcaaataag	aagagaaaga	tagagacaga	tgaaaaagtt	21720
aagcaaggat	gactatgatt	tcccatagtg	aatgtctgga	tgatgatata	attaaatgaa	21780
aattttttaa	aaggcagatg	gtagaaggag	atttgaagga	aagacaagaa	atttgtttgt	21840
tttggaattt	cttgtagaat	gacctgcaaa	attttgtcct	aataaatgtt	aacaagtggc	21900
tttccataca	aaaaccaa	caaacaaa	cctgatgtaa	accaaaaata	aaattctgag	21960
gccccctt	accatctgaa	tgaacttct	cctctgcaag	ggcactctta	aaatttaaca	22020
tgaaagactg	gttcagggtc	tgacgggaag	tgggggtcgg	acaggcctca	ttatgcctct	22080
ctggcattaa	catcaacaca	gaccttaagt	ctgttaagaa	gcatttacia	tctattctct	22140
ctgaagcctg	ctacctgaag	gcttctctctg	cacactgaga	actttggtct	ccacaatcct	22200
ttatcttaag	ccagacattt	cctttctatt	gatcccagg	ctttagataa	actcaaccaa	22260
ttgtcaacca	gaaaaatttt	aaatctatct	ataaccta	agccccact	tcaagttgcc	22320
ctgccttttt	gaactgaacc	aatgtatttc	ttaaacttat	ttgattgaag	tctcatatct	22380
ccctaaaaac	caagctgcac	cccaaccacc	ttgggtgcat	gttcttagga	tctcctgagg	22440
gctgtctcct	gagggccaag	atcactcata	tttggctcac	cataaatctc	taaatatttt	22500
acagagtttt	actcttttca	tcgacactga	tttatattgg	atttccaatg	gtgtaatat	22560
catgcggttg	gcgattttca	agctagctgg	tagagatatg	caaacaaaca	aacaaacaaa	22620
aagagatatg	caaaaaaac	tcttgtgagc	ttgtgtgctg	gctccaacac	accatgctgc	22680
tgatatttct	accaggcagc	atagggcgca	cggtgggtga	gcctggcaag	gagacaagag	22740
ctggaggtac	agggtctaaa	tgaagccaag	ggcgtgaata	atattgatga	ggagagcaag	22800
tagaaatgaa	aagagaaaac	cagtattagg	agatctataa	agaaacaata	gtcagaaaaa	22860

gagtgggttaa agaggcagaa gaatcagaag agaaccagat tcatgaaata aaagagagga 22920
 atcagtttca ggaaggaaaag tgtgggtcagt agtgtcaaata accasagact gaatagycta 22980
 aggattttaa agaggtcaac agcwcaccca attaagcggg ttttctgagc ttaagatatt 23040
 tttcttttatt tttctatggt aaaaatattt tataaagaaa aagaaaaaaa gagatcacat 23100
 ttgccccccc aaccctgccc tggcctctct ctctctaagg aagtcgttaa tgaatttggg 23160
 cagaatagat tcagagctgt gagagtgaata atttgactgc agtgggtgag gtgttaataa 23220
 aagggagggg gtagagatga aatctttcca gcagtttggc tcttaagggg atagcttgaa 23280
 ggaaagaggg ttggaagggt ggttttttaa aagaattatc actcactaat caactagaaa 23340
 tccagtggaat tatgcagtac ttgtctgtaa tccagcaggc taactttttt tttttttttt 23400
 ttttgagacg gagtctcgct ctgtcaccca ggctggagtg cagtggcatg atcttggtct 23460
 accgcaacct ctgcctccca gttcaagcaa ttctcttgcc tcagcctccc aagtagctgg 23520
 aactacaggc atgcaccacc atgcagtggc aaattttttt tgaattttta gtagagatga 23580
 gtttcacccat gttgggtcagg ctgggtctcaa attcctgacc tcaagtgatc acccgcctca 23640
 acctcccaaa gtgctgggat tacaggccag agccacagtg cccagccaag gctaacctct 23700
 tgatcccaat gacaaacaga acaaacatct tactcaagtc cagaagcaat aataattttg 23760
 aatcttgctt gcatgtcaac aggagccaca ttaatacaga agaggatcac attggtccaa 23820
 ttaaattgaa ttgattgaga gcctctgcaa tacacggctc actgcacaaa taatgatggg 23880
 tcctgggtaca tttttatttg accattgatt gctcgatttg tttctgtgtc taactgtgta 23940
 ttggaattaa gctgactcaa tttgaactgc aggtcctttt atccctctta ttttttattt 24000
 atttattttat tttattttta ggcattttct actctccaga aaatctctaa gatttcagct 24060
 attaggcatt tgtctttttt tttttttttt tttttttttt tgagactgag tcttgctggg 24120
 ttgccaggc tggaaatgcag tgtcacaaatc tcggtcctac acaacctcca cctccctgat 24180
 tcaagtgatt ctctgctc agactcccga gtagctgtga ctataggtgc cctccctgat 24240
 cgcccagcta atttttgat ttttagtaga gacagggttt cactatggtg gccaagctgg 24300
 tctcaaactc ctaacctcaa gtgatctgcc cgccttgccc tcccaaagtg ctgggattat 24360
 aggcagtagc caccgtgccc caccaggctc tcttaatcat aacaaattca tcttaaaaca 24420
 tcatttttaa atatattatt ttttattctt cattgacttc ctgcattact ctattttttc 24480
 agagtttcta gcaaccagta tcattggtat tttaaacaat gtgtacatgt acatttatgc 24540
 agatgagtta acatatatca aagcaacctc caaacaatgc catttaggta atctccaatt 24600
 taaaagcctc aatagaatga taagattgag cttttctgta gttccatgac ctccagcaga 24660
 gtctgcaagg ccacagctgc ctgaagggtt attctgtaat tagaagatgc cagggtcatc 24720
 tcagaataga acctcaagcc acccaggcta catttacaga atcagcctct ccagaaaaac 24780
 agcaacaaag gagggccttc ctatgtattt ggaaggagtc acctagagga gggacttggg 24840
 gttttggtgt tgggtgtggg ggcagggatg ggatggggag ggggaagctt attgaaatat 24900
 actaaaagac aaaccaacct aagggtcgga gggaagaaaa ttcacacttg taagcttctt 24960
 ttttaagggg catctcttag gctctagctt ttgagattca gtatatatat atttttgagt 25020
 cttgctctgc tggagtgcag tgggtgtgat tgggctcact gcaagttctg cctctcaggt 25080
 tcacaccatt cttctgcctc agcctcccga gcagctggga ctacaggcgc ctgccaccat 25140
 gccagctaa ttttttgat ttttagtaga gacggggttt caccctgtta gccaggatgg 25200
 tctcgatctc ctgatctcat gatccgcccg cctcgccctc ccaaaatgct gggattacag 25260
 gcgtagacca ccgcgcccg cccagtattt ttgttttatg aagatattac atttgtaaag 25320
 tatgagcttg gtgtcagcaa acttatatcc ctgtgtacaa actggccgaa tcacttagcc 25380
 actttggggc aaatcactta gctcttctaa cagtaagaaa tcaacaagaa aaataaacat 25440
 ttcaaacatt acaatgtgtt catgtattca ctgtggggga tgaccagatt ctcgaaacca 25500
 caggttgctt ttagtgaaac aagtttgggt tggggccata gacttggtga tttagaatca 25560
 atggctgttc tctctctggg actttgattt ttttcttggg ctcatccctt tttttagta 25620
 tcttatcttt gtcttatttg tataggactt aactgttccc attcccttat tagagcaatc 25680
 taagtgatta ccttcatacc ttttggaatt atatgcttca aaattccaaa agaattgatt 25740
 ttgggctggg cacagtggct cacacctata attccagcac tttgagaggc tgaggtggat 25800
 cgctgaggt caggagtttg asrccawcct ggccaacgat agtgaaacct cgtctctact 25860
 aaaaaatata aaaaattagc cgggcatggg ggcaggtgcc tgtaatctca gctactcggg 25920
 aggtggaggt tgcaatgagc ccagatcgca ccattgcact ccagcctggg caacaaaagg 25980
 tgaaactcca tctcaattaa aaaaaaata atgattttgg tggctgactt caaataggta 26040
 ggagaagaag gagagaggag atggagggtc asggagatct aattactctc taaaatcatg 26100

ctaggaaaga	taacaccttt	taataacact	ctctgctttt	ataacatcat	tctgccaaag	26160
agctcaaagg	tttcaacama	gttcactttc	agaaaacccc	tttgaggaag	acagaatata	26220
catctttctc	ccmttttaaa	gatgaagaaa	caggccgggc	acaatggcta	atgcctgtaa	26280
tcccagcact	ttkggaggct	gaggccasar	gatcgcttga	gctccaragt	ttgagaccag	26340
cctggataac	atggcaaaac	cctgtctcta	caaaaaaaat	acgaaaatta	gatgggtgtg	26400
gtggcatgca	cctgtggtcc	cagctacttg	ggaggctaag	gtgggaggat	cgcttgagcc	26460
caggagagtca	agtctacact	gagccatgat	tggatcactg	cactccagcc	tgggtagaca	26520
gagcaagacc	ctgtctcaaa	caaaatgaat	gaaagagaaa	gaaagaaaga	gtgagaggag	26580
aggagatgag	gggaggggag	ggtagcaggg	aggggggggag	gaaggaagga	aggaaggaag	26640
gaaaaaaaga	tgaaaaaaga	aatacgcaac	atgaaacaga	ggcagaaaga	ctttacgtaa	26700
attgtctcatc	atgtggttgt	caagtttgac	cccaaaaccc	aattttattga	ccaaggttat	26760
tctttgactg	aggcaagggg	gtccgctctc	ctgggccttg	ggcttttagaa	agctcatctc	26820
tggcctttct	gagatccatc	cctttctttt	tatttttctt	gacacggagt	cttgctctgt	26880
cactcaggct	ggagtgcagt	ggcatgatct	cgactcactg	taacctctgc	ctcccgggtt	26940
caagcgattc	tctgcctca	gcctcctgag	ataacaggcg	ctcgccacca	catctggcta	27000
atttttgtat	ttttagtaaa	gactgggttt	catcatgttg	gccaggttgg	tttcgaactc	27060
ctgacctgag	gtgagctgcc	caccttggcc	tcccaaagtg	ctgggattac	aggcatgagc	27120
cactgcgccc	agctcagatc	catccctttc	taagggcaaa	cagtccatgg	tgcaaagggg	27180
ccatgccacc	cagagttatg	agtacctggg	actccagaat	tccttgccctg	gtggcctcca	27240
catgcacttc	cagggcctgc	ttgggcctct	tctatgggtc	tgtcctgagt	gttgatagaa	27300
ccactgatgt	gagtacctgg	gcttgagccg	tggcctggag	atcctgttga	ctgtagcatg	27360
gagggggcct	gtgcagctga	atgtctgyat	gcagggtggtg	ggagttctgg	aatatgatgg	27420
agctggagggt	gggaagagaa	gtaggcttgg	ggcagctctc	tcatgccacc	tcattctggc	27480
caaaactcag	gtcaaactgt	gaagagtcta	aatgtgaatc	tgccttcaa	gggtgctaca	27540
aaggtatctt	tgtcaaggta	ggagaccttg	tggcctccac	gtgcacttcc	agggcctgct	27600
tgggcctctt	ctacgggtct	gtcctgagtc	ttctatgnaa	tctgtccttc	agggcagatt	27660
catatttaga	ctcttcacag	tttgacctga	gttttgccca	gaataagggtg	acatttagtt	27720
tgttggtctg	atggatgact	taaatattta	gacatatggt	gtgtaggcct	gcattcctac	27780
tcttgccctt	ttttttgccc	ctccagtgtt	ttgggtagtt	ttgctccctt	acagccaaag	27840
gcaaacagak	aagttggagg	tctggagtgg	ctacataatt	ttacacgact	gcaattctct	27900
ggctgcactt	cacaaatgta	tacaaactaa	atacaagtcc	tgtgttttta	tcacagggag	27960
gctgatcaat	ataatgaaat	taaaaggggg	ctggtccata	ttgttctgtg	tttttggttg	28020
tttgtttctt	ttnnntnnnt	nnntgttttt	gtggcctcct	tcctctcaat	ttatgaagag	28080
aagcagtaag	atgttcctct	cgggtcctct	gagggacctg	gggagctcag	gctgggaatc	28140
tccaaggcag	taggtcgcct	atcaaaaatc	aaagtccagg	tttgtggggg	gaaaaacaaa	28200
gcagcccatt	accagagga	ctgtccgcct	tcccctcacc	ccagcctagg	cctttgaaag	28260
gaaacaaaag	acaagacaaa	atgattggcg	tcttgaggga	gattcagcct	agagctctct	28320
ctcccccaat	ccctccctcc	ggctgaggaa	actaacaaag	gaaaaaaaaa	ttgcggaaag	28380
caggatttag	aggaagcaaa	ttccactggg	gcccttggtc	gccgggaacg	tggactagag	28440
agtctgcggc	gcagccccga	gcccagcgct	tccgcgcgct	cttaggccgg	cgggcccggg	28500
cgggggaagg	ggacgcagac	cgcgaccct	aagacacctg	ctgtaccctc	cannncnncc	28560
ccaccccacc	cacctcccc	caactcccta	gatgtgtcgt	gggcggctga	acgtcgcccg	28620
tttaaggggc	gggccccggc	tccacgtgct	ttctgctgag	tgactgaact	acataaacag	28680
aggccgggaa	cggggcgggg	aggagggaga	gcacaggctt	tgaccgatag	taacctctgc	28740
gctcgggtga	gccgaatcta	taaaagggaac	tagtcccggc	aaaaaccccg	taattgcgag	28800
cgagagtgag	tggggccggg	accgcagag	ccgagccgac	ccttctctcc	cgggctgcgg	28860
cagggcaggg	cggggagctc	cgcgaccaa	cagagccggt	tctcagggcg	ctttgtcctt	28920
tgttttttcc	ccggttctgt	tttctccctt	tctccggaag	gcttgtcaag	gggtaggaga	28980
aagagacgca	macacaaaag	tggaaaacag	gtaagaggct	ctccagtgcg	ttacttgggc	29040
gttattgttt	tgtttcgagg	ccaaggaggc	ttcgggaagt	gctcggtttc	ggggactttg	29100
atccggagcc	ccacatcccc	accacttgca	actcagatgg	gaccggaggc	ggtgttaaat	29160
ggggagacga	tgtcctagta	cgagctctgg	tgaccccagg	actctgcgct	gctgcgcttg	29220
ngggcttgcc	cgacgggtga	gaccggggag	catctctggg	cgtggagacc	cgggcgcagt	29280
accccgggct	cagaggggtc	gggggttccc	gngcgtgct	gagggcgctg	ctgccgggtg	29340

gggagagctg	caggtccggc	accgagncgc	tgcctttgttc	ggagggccct	gagctggcnt	29400
agnaaacct	tctggttgca	ggtcggccag	tacctacgga	gacaaatgcc	agcactgagt	29460
cttcactcgg	ttcttaagaa	gctgggtctgt	tctgacctgg	gaattggcta	tatgctcccc	29520
gggactggag	cggcacagtc	cgggactgtg	aatccgggaa	ctcagattgg	aggtgtccca	29580
aacggtccgt	ggtgctattg	ctcactagag	gccttgggtc	tttgnnttga	cctgaggggt	29640
agggaggtcc	tgccacagtc	ctccgtgcgc	tcagctgagc	tggtgtccct	ggcgagagc	29700
gcggacgagt	tttgtttcct	tttctttttc	ttttttttct	ttcctttaag	tctcggctctg	29760
tcgcccaggc	tggagtgcaa	tggaacgnat	ctccgctcac	tgcaacctcc	gcctcccggg	29820
ttcaagcgat	tctcctgcct	cagcctcctg	agtagctggg	attacaggcg	cgtcaccaca	29880
tccagctaata	ttttgtatct	ttagtagaga	cgggggttca	acttggtggc	caggctgggtc	29940
tcgaaccttc	gacctcaggt	gatccaccgn	gcctcggcct	ccccaagtgc	tgggattata	30000
ggcgtgagcc	accgcgcgcg	gccgagtttt	gtttctttta	aaaacaagac	ttaggagagc	30060
ctgcggagac	ccggaggtgg	ggtgcccatt	cctccctctc	ccacgttccc	tgcagcccca	30120
tcttcagac	cgttgctgct	ggtctctcgg	ggcagcttct	gcctggggcg	agatggggaa	30180
gctgggcca	ggtggtgccg	tggaatgacc	gggagtaacc	ccggcggggc	gcgcagaact	30240
cggagctccg	ccgcggggct	gggctgggct	ctgccgtgag	ggtgggggtg	ctgggcgcgc	30300
gggctgcggt	ggnccccgga	gactggcccc	cagncgcctc	ctggccggag	gacctaggaa	30360
tcggccggct	ctactagggt	tctttgctcg	cggttccgac	tgtgaatccg	gtgaagaccg	30420
gtggttgag	acggggagga	actatgaggt	tgaggcgaaa	gcccgttttg	tttttttttt	30480
tttggttttt	tgggtttttt	ttttgttagc	gtggttgcca	actcccaggc	catttggttaa	30540
gcaggaaggt	tcttggggcg	gcggacgggt	ccagggttat	gtgttaggtg	ctcttttagt	30600
atatctttta	tcaaaaagaa	gcaaaagaaat	aagattaaaa	ataaacaaaag	aaaaaagttg	30660
tctggcactg	gcagtaattg	gcctgccttt	gcagcactga	taccattagc	ttttaaaatc	30720
cgacttttca	ttgacacttc	aagaagagaa	tgggtagtat	atacacattc	atctcatagt	30780
ggacaaaatt	catatttaaa	aaaaccttct	gggtactgaa	atcagcaagt	cacttgccct	30840
ccatggccga	atccctgctt	cccacgaaga	gaacctcaca	aaaatttccc	ccaagttaaa	30900
gagtggaaat	ttcttgattt	ttttnttctt	ttttttttta	cggccgtagt	ttagaanccc	30960
agacttaaat	tatgatcttc	ttttcaaaaa	aaacttaaaag	tccttaagtt	ttcatctccc	31020
cttttatatt	aacctattct	tctcatacct	accacaaaaa	taatggaggc	tttctgttga	31080
gaaactttcc	gtttctgttg	agagtatcat	tctcttgaga	aactttctcc	taaatcagag	31140
aaagtatgga	agcatggaaa	gtattcctga	gtagaacctc	tacagatatt	acaatatatt	31200
tcaaatacaa	agtttccatt	gtcagcctgt	ttcccaagtg	cttcacacaa	ccattaaata	31260
attccacaaa	ccattaaaaat	aattaatgct	agggaaatct	aggaaaacat	tggttttacaa	31320
tcagaaggac	cggggaagtg	ggtcttcagc	cttcacgatg	actacaagcc	atttaaggga	31380
ctagaattgc	tactgttgct	agagcaattt	aggagtctgt	atgtgagcac	ccgcatagtg	31440
ttccagaatg	acatatctga	ctgtaacctg	gacacgtgtg	atatgttgct	tcccctgcag	31500
atgagcattt	gaaatctcaa	ccctcgtatt	tctacgagtg	caggcctata	atggaccctg	31560
ggcacatttt	tttttttttt	gagatgcagt	ctcgtctgtg	ttcccaggct	ggagtgcagt	31620
ggcacgatat	ggctcactgc	aacctccacc	tcctgggttc	aagagattct	cctgcctcag	31680
cctcctgagt	agctgggact	ataggcgcac	gccaccatgc	ctggctagtt	tttgcatctt	31740
cagtagagac	agagtttcac	catgttggcc	aggatggtct	cgatttcctg	acctcttgat	31800
ccacccgcct	cggcctccca	aagtgcctgg	attacaggcg	tgaggcaccg	cgcccgaccc	31860
ctggacacat	tttgacttag	aacatatatt	cggtttgtgt	gagacagtgc	attagtgcag	31920
gattggaaaa	gagtgatcag	gaattgattg	ttttcaagga	ttggttcctt	ctgctcaagg	31980
aagtccattt	gtaaacataa	aaaaatgaat	gaaactgaag	aagttcagtg	acttagcttt	32040
ttattatctt	ctgtagtact	tacctttttg	gagaggagtt	ggttgggata	tttttccatt	32100
taaaattttt	ttttaaaggg	atcttctctc	ccgtaagccg	ggatacttaa	gctatatatg	32160
tagtggctac	aaattaagggt	cttcactgtt	ttcattttta	gctgctagaa	taagtgaaca	32220
ttaccttaga	tagactcttc	taattatgaa	gatatctaga	tgtctagaaa	atatcaaaaat	32280
gcatgtggtt	tttgcatctc	taaaataact	ttaaaaccaa	atactttttc	tttttttttt	32340
ttttctgaga	tggagtcttg	ctcttttgcc	taggctggag	cgcagtagaa	tgatcttggc	32400
tactgcaac	ttccgcctct	caggttcagg	tgattctcct	gcctcaacct	cctgagtagc	32460
tgggattaca	ggtgcgtgcc	accacacccg	gctagttttt	gtgttttttag	tagggacagg	32520
gtttcaccat	gttggccagg	ctggtctcaa	gctcctgacc	tcaagtgate	tgccagcctc	32580

```

agcctcccaa agtgctggga ttacaggcat gagccaccac acctggcctc aaatacattt 32640
ttttaagtat ccagatatata aataaataat accattatag tagttgttat ggtcatttac 32700
tctagcatca aatgttataaa gatcattctg aacacttggt ttgtttatgc tgagagaagg 32760
cctactccaa aaaatgcaac catcttcgta tctgcatgtg gatacaacca tgaatggcca 32820
aagttattgc agtgttgaat agacacttat atagcactgt gtggcaagta ctgtttgaaa 32880
tgtttttcac gtgtttattc attgtattta ttttgagata gggctcttgc ctgttgcaca 32940
gggtggagtg cagctgcaca gacaaggctc actgcaacct cagcctcctg cgctcacgtg 33000
atcctcatac atcagcctct taaggagctg gggccacagg cacgcaccac tgctcctggc 33060
taaattttta caattttttg taaagacaag gtctcactat cttgctcaga ctggctctga 33120
actcctgggc tcaagtgatc ctcccacatc agcctcccaa agtgctggga ttacaggcat 33180
aagccactgt gcaggtcatc aaatgctaata tgaattttca caacaaaacc atttattgtc 33240
cctagtttac aagattaagt aaatgagaag ctaatttttc tctggctata taccttgcaa 33300
gaggcagagc taagacttga acccagccag agttctttta ctccagcact aacatttcag 33360
ctgctgcaac cagggaagctt ttcaaggatg atcaccacat tctctacatt catctgctat 33420
aatccttatc agaacttaca gcctgtatca tattttcctt gttgctgtga gtgggtcagc 33480
caaattctct ttaacttgaa accttggttg cgtagggatt gcaacatcct ggaaagaata 33540
gaataaaatt tactcaactc aattttttac ttggttcata atgaaaacta tactattgct 33600
tcagtcagat gtttgcgaaat agctgtgtga tctcaaaatg ttttcctatg tgatctatag 33660
naaaatggaa tgatagagta ttaggctgta agggcctaag anaacaaagg aaaaagagaa 33720
gtgaactgtt agtttagttg taaaacctaa ctttggtgaa ttgtaaaaat ttgttataat 33780
acaatatgat tcttgcttgt cctgtccttg atgaagttgt ggaccttttg aaataagcta 33840
tttctctgtt actgctgtta ctgttttagaa atcaaattta gtttttctt aagatatacg 33900
tatttttgga agataaacac agtttcaaag tctgccttgt tggctgggtg cactggctca 33960
ttgttgattt cccagcactt tgggaggcca aagcaggagg atcacttgag gntcaggtag 34020
ttncagacn cagncctggc aaatatgggt aaaccccgtc tctactaaca atacaaaaat 34080
tagctgggag tggtgggtgg tgcttgtaat cccagctact gggattggga ggctgaagta 34140
gaagaattgc ttgaacctgg gaggcggagg ttgcactgag tcgagatcgt gccactttac 34200
tccaacctgg gcgacagagt gagactccgt cttgaaaaaa aatgtctgcc ttgtaaaagt 34260
gaaataggat gagaaagtgc ctttcttatt aattggtgta atgaattaga aataaactct 34320
ttgaagacac ctcttggtta aaatagttac atttactgtt gatttatggt atgttggata 34380
tgtttttaag ttttcctgtt aataactcag ttcattctca tgagtgaat aggtgctttt 34440
attgtcttta tagatgggaa actgaggtat aggcaggcta ggtacattat tatggagttc 34500
gtaagtagtg gagctgaagt cgcacccag acagtttggc ttccgtgagt ttaccaatct 34560
catggtaaag actttgtcag actatcaaag ttttgacaaa tgaaatatta gcaaaaggcc 34620
aaaagggtt ctctattttc atttgagtat cttcacctga aaatagttgc ctgaataagt 34680
agcctgcata gaaaggtaca ttttagaaat acttgaggcc agagaatgaa aagcttacat 34740
aaaattgatt tccggtgggg ccttcagtta ctctccattc tacgaagacc acaaatagca 34800
ttcaggcaaa gagcatttat tccaacaatg gaggagcact ggatttggtt cctaaaacaa 34860
aataaagttt gaaatcctgt ctttcccatg ttgaaaacaa agttggtaca aaacctttt 34920
gcttttgcaa acctccttta agaccgatt taaatgcytc cctcctcatg aagctcttct 34980
ggatccactc yttcccatca ctaagttgaa agtaagatcc ccttctcttt acttccatta 35040
gacttggtt acagcactct ttgtatcatg tatttaattc tgttttttaa ttacagttaa 35100
catttatttg tcttctctct gagtgtatgc ttctctagag gaaggtcttt gattcattct 35160
cccttgacct taattcatcc cacttaatat ggaaaaaatt taataaatgc tgacttgaat 35220
aagtccaaca aggagaatgg caagctcatg tttgcttctt gtcttctaaa agactactta 35280
agataacagg gtaatcacag aaaagcatta gaaatagagt tatatgagaa acaactgtag 35340
ttaaggctag gtttatgtta gactgagaaa ttttagtgca tacttaagtt atttaggcca 35400
ggttactttt tgtagaacia acatttcagt ttcgctcagt ttcatttccg tttctgnagg 35460
cagctgtgat ttaagaaaaat gctctagtct gtggcattcc atattcaagt actttgagtt 35520
gtatattaat ttatttttgt taataagagt gacatgactc actaagtaat ttagagattt 35580
aaacactttt ttaaaaaaca gtaacttcat atgcattgga tctattcttc tataaagtct 35640
tttcttgggg ggtgtttgtt taaaattccc cgggtgtttt tctgccaaat ccaacttcca 35700
agaagcattt ggaagtcaaa acattttatc tgggtagtct taaagtccag atattttgtg 35760
atagctggta tttagtttat gatattttccc aggaagaact ttttagtagt tgaaccattt 35820

```

atgaaagact	tccttgaagc	taccttagag	agttgattta	gttcttcccta	aataagtaaa	35880
tagaatatta	gtattaggac	atcttggagt	atagatgcaa	atattggtga	aaaagaacat	35940
ggatatcaga	gtcaaattaa	tgtagattgg	aattctggtt	attactaatg	gatatctgac	36000
attaggcaag	ttgctgatca	ctctttgcct	cagtttcatc	atctgtaaaa	taggtattttg	36060
tgtttgtgta	taatgtgaac	cgtataatat	aatgcttggc	ctatagtga	atttatttcat	36120
acgagtgttt	tcagtgattt	taaaagcttg	ctttaggccg	ggcgcgatgg	cttctgccta	36180
taattccagc	actttgggag	gccaaggtgg	gcgcatcatg	aggtcaggag	ttcgagacca	36240
gcctgaccag	catggtgaaa	ccccgtctct	actaaaaata	caaaaattag	ctgggcgtgg	36300
tggtgcaagc	ctgtaatccc	agctactcag	gaggcttagg	cagaagaatc	gcttgaaccc	36360
aggaggcaga	gattgcagtg	agccgagatg	gtgccactgc	acagagcgag	actccatctc	36420
aaaaaaacaa	aacaaaacaa	aacaaaaatc	ttgctttata	gtttacttcc	acatcaaatt	36480
gtctttatcc	catgttactt	gcattgatat	cccagacatg	aaaagaaaaa	aagatgataa	36540
caatgacagt	tattaaatta	ggttccactc	ttattctaga	tcaccaattc	atattactat	36600
tcagacttgg	aacattaaat	tttagttaaa	ctttttttca	aatatgcata	taattgtcag	36660
tggttactat	attttgggga	agagattggt	gacttctttg	aagaaagata	cggatttttct	36720
cttcagaara	aatacacatg	ggctcatata	atccaaattt	tatgtgtaat	tacagggtgt	36780
tcatgaatgc	ccacaaatcc	attaagccat	gtgaccttgg	acaagtcatt	ttacttttct	36840
gttttttagg	ttgttgggtc	gtaaaatgat	actacttgac	ttttaagag	cccttcaagc	36900
tcttatgtcc	tctaacccca	ggctctgtatt	cagaagaagg	ggtgggcctt	taattagagc	36960
catctagaga	tctgaggaac	atgctgggca	ttagtgtaac	ataccatgtg	gatttttgaga	37020
ggtaaaagaa	aaataaccag	ggaatgcctc	agagcattcc	tgatcagatc	gatgacagaa	37080
gaaaggaatg	agagggagga	gaggaagctg	ttgaaatttc	ctatttacct	gctttgagtg	37140
aatgaagatt	tgaatcatag	aaccagaagg	ggttctcatc	tgaaatgcaa	aggaaggagg	37200
agttggttta	attcaataag	tttcagttga	gtaaacatga	tttagtgaga	tactgttctt	37260
gcttctgact	caccattttg	aaaatctctc	taaaataaaa	ttggactctc	catctcggac	37320
atcatttttg	gtgtagggtt	tgcttttttt	tttgagatgg	agtctcgtta	tgttgccag	37380
gctggagtgc	agtggcgcaa	tctcggctca	ctgcaacttc	cgctcccag	attcaagcag	37440
ttctctgcc	tcaacctctc	gagtagctgg	gactacaggc	gtatgccacn	catgtccggc	37500
taatttttgt	atttttttta	tagagacggg	gttttactat	ggtggctagg	ctgggtcttta	37560
actcctgacc	ttgtgatctg	cccaccttgg	cttcccagag	tgctgggatt	acagatgtga	37620
gccacagtgc	ccggcctaag	ttttacttct	tataatggac	tctgtttaag	ccaatagggtg	37680
atgaaaggaa	accataacca	actcttcagg	ctcattcatc	cttcaagaat	agcatgctag	37740
taccatcct	aggaggagaa	ttggactata	cctcatgagg	atagtttgaa	gtatctcaga	37800
agaccctcac	tgggggtagg	tccgtaagac	acaaagcttt	ctaaagcact	gtaccaaatt	37860
tgttgtttga	gagatcataa	caaattagaa	gtggaaagaa	gaaggagtaa	aaggaagaag	37920
aggtttcttg	ccaggagcag	ggagggggaa	ggagctgcta	ggaagatgtt	tggttgtcat	37980
atccctgttc	acccttgctt	tgcaaaattc	ttgtaggatg	ccaggttggg	agtaattgtt	38040
tttcacaaga	gtcaaaccac	gcttgttttc	ttgaagaagc	aagtcttttg	agggtggtgg	38100
cttgaaatct	gttggatctg	gtatttaggt	gatacacttt	gacataaagg	ncaacactga	38160
tgcaagcagc	agctttccct	ggaaaggcag	ggagaaagtg	aaggcccaga	ctgatgagct	38220
tacactgacc	tgcagaccct	tctccattcc	caggcatgtt	tgggtggcaga	gtttacctta	38280
gttgggggta	ggctgttgct	tggtagctgt	agagagaagg	aagaagaaga	tatgatatta	38340
acaacaacaa	tacatattta	tatttgaaaa	ttaaatgcac	taatacacct	atagtgtctat	38400
taaacaatt	tatttgttca	gcaaattgtt	gttaaacacc	aatgtactgt	ggaaagtact	38460
aggtgtcga	cagaggtcaa	aagacttttt	aagaatctgc	caccattaat	gatctctttc	38520
tgcttggcat	tcaaggctct	ttgaaataag	actgtgaccc	actttgatag	ttttgtcctg	38580
gattataaga	cacatgctcg	aaggaactat	agctgggttt	ctcaccagac	tgattaacat	38640
atagtatggt	ttggtacctg	ttaaatgagt	ctctctctac	aggttttcat	cttctacttt	38700
aagaacccct	tcctggcatt	gtttggctcc	tcattgttct	ggaatctcat	gtccatctca	38760
tgtatttgcc	ttggttcaca	cttgatcttc	tgctcatatt	cctcacctac	tgaaatttta	38820
cccatcaacc	agaccgtggg	tgatggaaca	cagcatgggc	tagttcttct	tatatatgat	38880
ctcatttaat	tttactgga	actctgagat	aggtagcatt	tcagcccact	caagttgact	38940
caaatttttg	taatcatcaa	tatattttaa	aataacttta	tatttgacct	gtaattggaa	39000
aaccaatatg	agttatcata	aatgaaaggt	aattttaaaa	ataattagga	tgaagacaaa	39060

ttatTTTTct	cacagtctat	gtataagata	aactattggt	tcccaaaggc	cccagctgac	39120
aatgagactt	ctcttacttt	gttgaaaagg	gaattagcaa	gcattaaaga	ggtgtcaaaa	39180
agaagactaa	acaaaagctt	actccttttt	ttttttgaga	cgagtcttgc	tctgttgccc	39240
aggctggagt	gcagtggcac	cacctcggt	cactgcaacc	tcttcctcct	gggttcaagc	39300
gattctcctg	cctcagcctc	cctagtagct	gggattacag	gtgcatgcc	ccacaccg	39360
ctaatttttg	tatttttagc	agagatgggg	ttcgtcatgt	tggccaggct	ggtctccaac	39420
tctgacctc	agggtgatcca	cccacctcgg	cctcccaaag	tgctagaatc	acagggtgtga	39480
gccgccgcac	cgggccgctt	actccttaat	gtactaagaa	tgttatatat	aggctgaaga	39540
agtgtgaaa	agaaccatat	tttctcatga	tgtggttcaa	tgtttaatac	tgtgcttggt	39600
catctcctaa	aatcctctga	atatcactta	aattcatcct	gtgtaactct	cccacatttg	39660
gggaaatact	gagcttgctt	attattatat	tagccccata	tttcagatga	tgcacctgag	39720
ccgaggagaa	gttaaataac	ttgttcatgg	ttccatattt	ggctaattgg	agagccagg	39780
attcaaactc	ttgtctctct	gactcccagg	tttgtgcttt	tcccacttgg	ctgaatttct	39840
catgctacct	cctccataac	accttttcta	gaacttttaa	ggaatgcttc	ccctgttctc	39900
tcatagcatt	ttaatgtaa	ttgccaaa	gtcccagttt	gcacccctac	cagcaatgtg	39960
ggagaaaata	gcaacataat	tttgatgttg	gggtctagtg	ttacggtttc	ttctgctctt	40020
tgggatgtat	attccatggc	tactgaatac	ccaagtccca	aacagttttc	ttagactcag	40080
aaggctcatcc	acttcagctt	cttcatcaaa	aagacatatt	ttctggctgg	gcacgggtggc	40140
tcacacctgt	aatcacagca	ctttgggagg	tggaggcagg	agaattgctt	gaacccagga	40200
ggctgaagtt	gcagtggacc	cagatcgtgc	cactgccctc	cagcctgggt	gacagagcga	40260
aactctgtgt	caagaaacaa	aaaaaaggac	ttgttttctg	ttccattacc	cacagtggta	40320
gaatggcgctg	ctaaatttat	tctccagctg	ccattaactg	caaattaaaa	tcttagtctc	40380
ttgcctcttt	aatccaggct	tcttcatact	ataccagaat	ttaggataac	tattacagt	40440
ccctttatag	gagagaaa	agaaattgtg	tctgtagatg	tctgttctct	tcagcttaaa	40500
atggacactg	aaatgttaaa	tattggactg	gcctcattta	ttctcctgt	ctgttgggtcc	40560
aatttgaatc	ttaaggcgct	tttcaactgg	aattttttgt	ttctctcaac	taaaaattgt	40620
tctttgtaag	tttgaaatcag	aacaaaatcc	tgaatgttga	gggtttccta	aaggctgttt	40680
ctttatgcaa	aagcctgaaa	cccgatgttg	atgttggctg	cttaaaatta	actgtgaatc	40740
aaggcagggt	ttttattttt	attttttttt	tactttaatg	attgtgttaa	ttatagtga	40800
aaccttgagt	tcacgagaaa	gaaagccttt	ggtcaagtat	tgtttattaa	gttgtcagtc	40860
ttgttgagg	atgtgcaaat	ttagtggaat	tagtgccatt	tttcagttta	caattccagt	40920
cacatttcac	atgatcagag	catggctttc	ttctctgtgg	agcaaataga	gggctgtctg	40980
acacttgggt	ccagtggctt	ccattaagca	gagtggatat	gtccctggag	tctgcagaga	41040
agggcatggc	actctgacct	cagatggcac	tccgttttgg	gacattgtcc	aattctagtt	41100
catagcatat	gtgaccaaca	ccagctctca	cctgatgtaa	acacttagcg	cgttggttgc	41160
tgggggattg	gattgtgtga	atttttcaaa	actacagttg	acagaaggag	gctacccaaa	41220
atgaaaccca	ataattccat	ttttgggaat	tattcccact	tttgttccat	ttttcccact	41280
ttgttctttg	gcacacagaa	tgtttgattt	gtgaaaatct	taataacagt	agttttttct	41340
ataaggaaca	ctcagaatct	tgataatatt	ggaataatac	agatcctttt	gtaggatcct	41400
ctcagacctc	atataataga	gttcatgtag	tcaatattta	aagaaaaaca	cccttaagtt	41460
tttgtttttc	agaatcacia	gtaagtggat	ttaaacttgt	gatcttattc	ccctttcttc	41520
tcttaattta	gtgaggcagc	cagcgagagg	gtttgttttg	gttattctaa	agaaggagtt	41580
tgttgtaag	ttttggagg	caagacttag	actctgtgtc	tctgtgcttg	ccctggaaact	41640
ttgattaaat	tgtcactaac	cgagttagct	ggccctcgcc	gggctgcaga	aatagaagtg	41700
tcttgccac	atgacatatg	actgtctcaa	gagctggctg	gtgaaaggac	gttctggaga	41760
aggctgccga	tactgtatga	actagaant	ggacaagagc	ctggagattg	gataactcag	41820
tttgccgcaa	gtaaagggaa	taaaagtgtt	aagggtggcaa	aattgtatcc	aggtgtttat	41880
aggctccctg	agttcctgac	ttgagcctat	ctatgggttt	agagttcaag	gctctttacc	41940
agtgttgaca	atcttatact	ctagggttga	cctccgggga	aggtgccctt	gcttgatggc	42000
atgtttacca	gggtttctag	agcctcaatc	acagattctc	tctagctcac	atgaagttaa	42060
tgaaaatgaa	tgtgcttccc	tacaaattag	agaggctttg	aggaaaaatc	agattaaatg	42120
cactcctgct	tgaacttatg	tttcttagaa	cacagctgga	aattttgtca	cacaaacctt	42180
tactttcagt	gacatttctt	gactgggttg	ttactgtagt	gaatctgctt	taactatctt	42240
ttcttatcgc	tgagggtttta	cttccattct	acatgtgatt	gtggagcgct	gcgtcattgt	42300

gggttcagtg tagtggagag taggaagatg gtgagacaca gtagcttggt gcacattgct 42360
 taatattatca gggatcactg atgagttagt acactagaga agattgtagg tagagctgaa 42420
 aagatggagg aattataagg ctcagatttc tctctttttt ttttttttta agatggagtt 42480
 tcaactcttg tgtccaggct ggaattcaat ggcatgatct cggctcactg caacctctgc 42540
 ctcccggggt caagtggagc ttgatggct cacttgatgg tttcctggct cagcctcctg 42600
 agtagctgag attacaggca cccactacca tgcccagcta cttttttgta tttttagtgg 42660
 agatgggggt ttatcatggt gaccaggctg gtctcgaact tctgacctca ggtgatcacc 42720
 tgcctcagcc tcccaaagtg cagggattac aggcagagc cactgcgtct nggccaaggc 42780
 tcagatttct aatagagatn ttctaattgga catagaggct ggaggaaatg ggatnggaca 42840
 ngganaaaact gagtncnagg tgccaaannn aacttgtagg gggccgggtg cgggtggctca 42900
 nncgnnccnt gtaatcccag cactttggga ggctgaggcg ggtggatcac gaggtcagga 42960
 gatcgagacc atcctggcta acacggcgaa acccctctc tactaaaaat acaaaaaatt 43020
 agccgggctg gttggcaggc ggctgtcggg agtcccagct actcgggagg ctgaggcagg 43080
 agaatggcgt gaagccggga ggccggagctt gcagtggagc gagatcgcg cactgcactc 43140
 cagcctgggc gacagagcca gactccatct aaataaataa atagataaat aaataaaata 43200
 aaataaaaaac ttgtaggag gtggcagtg gctatggagg ataggtgcaa cctctgtgag 43260
 aatgtagaga aaatagtata agtgagtggt gaggaccccc aagagggggt tttatagtaa 43320
 aacaatggtc agaagtggca acaggatacc gtataatgct ttcacctcta ccaatgcact 43380
 gggtagctgga gagcgtccca gtttgctctg gaaaggccct ttctgtggac aaagaataca 43440
 gaaaagagat tcctttaata aaccacccac tctgtgtccc acccttgatg aatacttcac 43500
 tgtgaaattg ccagaattaa tcatggtaat agctactgta cacttacttt gttccaggaa 43560
 ctggataaat gttttacata cattatcagt tctatttttt ggagaagata caggggctca 43620
 gagtccctag gggtccagag ctggtgagta gcagagtcag gattcaaacc cagctttctc 43680
 tgactctaaa acctcctttc ttctgtctga aacaattaac tcaagacaac aaaggagtta 43740
 aggatttggg gagttttctg catggtagaa tagaccctaaa ggaaaagaaa gaaagacagt 43800
 gactaagatt tgggtttgtc tgccccacca aatgctttga gcactttcat aaatataaat 43860
 ccttcagggtt gggagaagg tgaacatctg aagacactga ttcttcagag atgtaatcca 43920
 aacaaagtga cttttggtga tatggtcact aaaccattta tcccaaaatt ctcttgaaa 43980
 accgtcctat aacagcaggg aacattatcc agccaagttt ttctgcaaat aaagggttgc 44040
 tgatagaggc ttgcctgctt gtgtttctng tnagcntnca ggggtgtttat gaattcacta 44100
 atcccttccc ttcagatccc ttttattctg gtgttatgat tgtgactgna aaaaaattga 44160
 ttttttttct atgacataga atgttgaaag gttgatttct tttctagagg aaagattctt 44220
 tttttctatg tgctacatac cccccgacca gggaaaaggc aaatagtggg attgtttgct 44280
 gaagtcttcc tttgaagggt gcttggtggt tgcttagtgg aaatcagcag ggggaagagag 44340
 gctatctctn aacattttgt tagangtttc ttctgnagtt ctatagtgat gaaacaagga 44400
 cttgggggtac aggacagatc tgctttcaga aatcctggct cttgtgaggt ttagaagccc 44460
 tgagaccatt tagctggtgg caacgggaat gttgaggggt ataaatagga tctttgggtg 44520
 tccaagtatc agtgacatga tatagatgga gttaaaccct taggatctcc ttatttattt 44580
 gtttgtttat ttttgagaca gggctcttgca ctgtagccca ggttgagta tgggtggcatg 44640
 atcataactc actgcagcct caaactcctg ggctcaagcg atcctgctgc ctcagcctcc 44700
 caagggtgta ggattacagg catgagccgc cacaccgggt caggatctcc tgtaaaatta 44760
 tattgttgac aacatgaaga attatgcttc tcaaaagcta gttatagatt tgtacaatat 44820
 tcatagattt cttgtttcag tttttacaaa ttcatagccc ttattttgaa aattagctat 44880
 tagcaataat tttgtctagg aaattggatg tgtattcaag tgaaagaagg aagtacagtt 44940
 acctattatc ttattgtaac taacaatcaa gtaagtgtga tgcatttggt actttaaaaa 45000
 ctgcacccaa gttacagatt attggaatta ataaaattca ctggatctat atatttttaa 45060
 acggacagtg tgatagcaga acctcttata gaatngatag aattcctctg gaatgattgg 45120
 ataacttcat ttcactcttg acttttacct tggaggattt cttacccctt ttggcttctc 45180
 aaatttgact attaaaatgt tgccctttaa aataggaaca cagtttcagg ggggagtacc 45240
 agcccatgac ccttctgcaa ggccccctaa ctcaaggtag tttccctgga actgtggttt 45300
 atggaatgtt tcaggagtgt gaggaggtat aatttaaggc tgtcctagca aggataccct 45360
 taaggataga gggccagta gcatctggag gccagaaaag ttanaactga ggcagtcaga 45420
 ttagcttcan ggctcaatta agctgatggg tcagcctggg agaaattgca ggatgactct 45480
 caatatcccc tcccaccccc acagcagcca cgatctgtct gtctttaatc atgggtgcag 45540

tgaacctgtt	ctttccaggt	gtcttggcct	tcagtaacct	tgttaggctt	gtccctgaac	45600
gtggctaccg	atccaaagac	acatgatcag	agaggcaatt	agagaacaga	ccttttccaa	45660
agcaagcatg	ttctgttggg	cttagaagtt	tcatgtccta	atattatagg	accctgtgca	45720
tctctctgga	gatgaggcac	atgagtcata	tctgtgattc	ttgcttttgt	gtcaacatct	45780
catgaatagg	caatcagagc	tttggcacca	atgtattttc	agttcatatc	tgatgtagtt	45840
aaatccacct	cctgctttgt	agtttactgg	caagctgttt	ttgatataag	acatctagaa	45900
cactgtaaat	atataacatt	tttatttgtc	tattatacct	caattacgaa	aaagacatct	45960
agaagcaacc	tcatacaagag	agatactgag	gccgggcatg	gtagctcaca	cttgcaatcc	46020
cattactttg	ggaggctgag	gcaggtagat	cacttgaggt	caagagtttg	aaaccagcct	46080
ggccaacatg	ttgaaaacct	gtctctatta	aaaatacaaa	aaagtttagct	gggcttggtg	46140
gtgggacact	gtaatcccag	ctactccgga	ggctgaggca	ggagaatcac	ttgaacctgg	46200
gaggcagagg	ttgcagtgag	ctgagatcac	accactgcac	tccaacctgg	gcaccagagt	46260
gagattacat	ctaaaaaata	aaataaagta	ataaaaaaga	gagatattga	tagctgttgt	46320
tggaaatttc	aacttccatc	tcacttctgg	taactttttg	gaagtttgtt	gaacaaagtg	46380
gaatacacgc	acatacacac	acacacatac	tctcttgttt	gtttaagggt	taatgaaata	46440
gctgtcatat	aatcactgtt	tttgaaagag	gagaattagt	tgctatctgt	acattttggg	46500
tatgtgaact	atttggtatg	aactctgaga	aatgcattca	gaacaacaaa	caaaatcata	46560
ggagaaatag	ctaagtggga	aggggcatat	aagagttgtt	gaaaaagtta	tttcttgaga	46620
aaccagctct	aatgctaggc	aagtcacttg	ctttggggga	ggcctcagct	tctctgtcta	46680
taagattgca	gcaggggtgt	agtgggaatg	agtcttcaac	attccaagag	attttatcta	46740
ctaatacgac	agtcaaatgg	agcatgactt	tgtggaagcc	tctcctcttc	caccagagg	46800
ggccaatttc	tctgtcccag	tgagatgttg	acacttgat	gatccctgct	tggagacttc	46860
cctcttctgg	aacctgccct	ggctcaggca	tgagggtctga	ctgtcaccct	tcgataggag	46920
cccagcacta	aagctcatgt	gttggcagtg	ttcttgcggg	aaggaaaaag	accagccagc	46980
ccatttggtt	ctgcacaagc	aaacagcttc	tggtagctgt	acagatacat	gcactttctt	47040
tcctcactgt	gtttccatag	acagatttag	tgctgtagaa	gagtagagg	cagtcacggg	47100
aaggagtgct	tgtttttctt	ttggctatgc	caaagtggga	aaaatcctcc	tatcttgtct	47160
ttttagtgtc	atcctctctc	cccttttctt	cttctttata	attctcatct	ctcatctctc	47220
ctggaaatgt	gcatgtcaag	ttcaaaagg	cacaatgttt	tggtagagg	gagggtggg	47280
aacacgtgcc	aggtgctaac	tagggctcatc	atttccccct	tcacagccag	cttcctgtga	47340
atgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtat	ttcttttgcc	47400
agcatcactg	aatctgtctg	ctgtctggta	ttccagggtt	tgggttaggg	aaaagtaaaa	47460
gtaattttat	aatcccagct	gtcatttaag	ccaccccntt	tgtgggtagc	atatggtcca	47520
ctctctcagt	tcattgtcct	aaagatgctt	catcagaaag	gaataacttc	caccccggtt	47580
ctctctgtcc	ccttactctg	ctttattttt	cttcgtcaat	cctaccacca	ccacccactg	47640
tttgaacaac	ccactattat	ttgtctgttt	cccatccctg	gtagaatagg	agccccatga	47700
atgaaggaa	tttgcttctg	ttgttcacca	ctgaatctct	aaggatatga	acacacctgg	47760
catgtgatag	gcactcgata	aatatttgtt	gtggctcatg	ggcaccttgc	agagttaagg	47820
ctgcagttgt	ttgtggaatt	tataagtgg	aatgaatatt	tatctactat	tcctcttcca	47880
aggcgatcac	acaataatca	ggctttacac	tatccagttc	ttaggtcttc	caagttatga	47940
cttgtgaggt	atgttaatta	tgataataga	aggcagttta	tttggttcag	atttattgat	48000
gtgtaattta	ccacagtaag	acttcccctt	tacaaaagta	tgatgagttt	tgacaaatgg	48060
atacacatgt	gtatctacca	ctgccatgct	ccttttcagt	ctgtcgtccc	ctccacccat	48120
gaccactgg	caccactgca	gtgatttctg	tccccttcat	ttcacctttt	ccagaatgtc	48180
atataaatgg	aatcatgcag	tatgtagttt	tttgtgtctg	gcttattttt	cttagcatta	48240
ggcttttggg	attcatccag	gttgtcgc	gtaacagtag	cttattcctt	tttatggctg	48300
agtaagtgtc	ccagttttat	ttatatattt	atttatgagg	aggtgtctca	ctctgtcacc	48360
caggctggag	tgcggtagcg	cgatctcagc	tactgcaac	ctccgcctcc	caggttcaag	48420
caattctcct	gcctcctgag	tagctgggat	tacaggcacc	caccgccacg	cccaactaat	48480
ttttatatatt	ttagtagaga	tggggtttca	ccatgttggc	caggctgatc	tcaaactcct	48540
gacctcaggt	gatccgcccc	cctctggctc	ccaaagtgtc	aggattacag	gcatgagcca	48600
ctgtgcccag	ccccagtttt	atttattcac	cagttgatgg	tcttttcgac	aactaattgt	48660
ttccagtttt	tggctattct	gtataaggct	tctataaata	ttcacaaata	cctaggatgg	48720
gatgactggg	tcatataata	gtactgtata	accttagcag	aaactgtcaa	actattttcc	48780

aaagtggctc	ttccatttta	caattccaca	gtgtattgag	tcccagtgct	tccatacaca	48840
tgctagcact	tttaatat	aatttagtgg	gtatgtaatg	atatctcatt	gtgggtttta	48900
tttgcaatttc	tctgcagcta	atgatgagtg	tttctgctta	tttgggaagg	ttttaattta	48960
gcagtctggt	gtattctgta	gatattaata	acttcaaaat	atcagtggca	tttgcagtta	49020
aaatttcctt	aaaaaattgg	ccaaagggtt	ccagcagtc	cttctgccat	gcccactg	49080
tatgaaacaa	ggctgaggtg	tggagattgt	cacatttttg	caaggagtga	tccacttggg	49140
tgactgatga	gaccagaga	gcgtacgcct	cgggcttgag	ggtgaggacg	ggcgggaagt	49200
cgactgcatg	gccctgctgg	ccttgggagg	ctgcccagtc	cttagctaaa	gctggcagtt	49260
atgggaaaca	gacttagatt	ctattacgtt	tttcaggatg	tcccaggagt	cacctgggaa	49320
gctcagcagt	cctttgtgac	tttcaagcat	atggtagaag	ctgctgaaca	cagagctccc	49380
tctttgggga	taatttgccc	aatcatttta	atcaggcttg	agaaatgagt	taccacaggt	49440
ccaggagtgc	tgccaccctt	gaattctgac	accctatttc	tcctatccgt	ctcttaatta	49500
attaagcaga	catccccaag	tgcttacgac	aagccaggac	ccttttgcat	actaaggaaa	49560
acagggatga	aggaaacaga	aatggtctct	gctctgactc	agaaggtaga	aatcctcttt	49620
cccagccaag	tcttcctagg	gagcacgtag	gaagggtctc	gaaccacagt	gtcagttgca	49680
ggggaggata	tcaggaaaagg	acattgaaga	agtggagacc	taagtttgag	acctaggcat	49740
tagccaggct	agcagtgcct	gaaaaagtgt	cttaggacaa	gagaactcac	cagtgaagtc	49800
ccagtggtag	gagagcgtgc	agcatattct	gagcctgtat	acacatctcc	agggcattgc	49860
ttagcagggtg	gggagtggca	agagagtagg	ctggagtcac	agaagggagg	ccaggtagac	49920
cttggtgagc	actggactct	atgttcagggt	gctgaggagc	tggaacaaagg	ttttaagtcg	49980
gggagaggca	tgttcagata	tttgggtctag	ctgagtaact	ttgggtgctc	tgtgacaaat	50040
ggttgggaga	ccagtgaggt	ggcagttgcg	gtcatctagg	agcaggatca	gagtggccta	50100
ttgactggga	tgactgtgaa	gtgggatcct	ttccagccag	taactggaaa	tgtgtatgag	50160
ggcagaagtg	agtgtactgc	atttgaacaa	ttgagaaatc	tagtacatag	tactgtctct	50220
tttatatctt	tttttttttt	ttttttgatt	ttggtttgtt	tggttactaa	cttggaacac	50280
tgatgtggaa	atgtcccttt	ggcttcagtt	acctgagcag	aaggggccgg	gcattgccaa	50340
actctcctct	taggacagaa	ttgctcccag	tattgatcat	tgtgttctga	ggtgggggag	50400
caaattgtgc	aggaggccag	gtcagtgcca	aggtgggtgg	gaggaattgg	agcaggaagc	50460
ttgcctaagt	gtgcccagca	aagccacgggt	agaactttct	actgtggctc	tatgctactt	50520
cttagcaacc	ttctccatgt	gcttctctgga	gagtccttgg	agtcagaacc	ttttcttga	50580
aaccagaca	ctttacttcc	aagaaaatgc	tgtccaagaa	aactcatcct	tcccttcttc	50640
tcatgaacgt	tgtgtagagg	tgtgtcttct	cttcttttga	gcttttccac	tcagggttta	50700
ggggaggtga	tattctatat	ttgggtttgg	ctctgggtac	tgcaacacta	ggctattaag	50760
atttcatcct	tactgctttg	cccctcctat	ctttccagaa	accacacatg	gatttgctag	50820
aaataatgga	acgtcctggt	tggacaggat	ataaccattt	ctcagctaga	ggatattgtt	50880
ggaatgaaga	aagataaatg	gggagaaggg	aactcacatt	gctttggcac	ttaaattaag	50940
ccatgtactg	tggtgggaaa	ttatttatat	tatctcgttg	aatccacagt	agaacacagt	51000
tgaacaccat	acaaggtaag	tattgtcatc	cttattttac	catgaggaaa	ttgatgctta	51060
gagagcataa	agccttggcc	aggggcacat	agttgggaag	ccggggctaa	ttcatgcctg	51120
ggctctttct	gatagttttc	cttttttaat	tgtccctccc	tcattgttac	cttggggatt	51180
tcaagagatt	catgtagctt	ctaaatcaac	gaactgattc	ctggagagca	gcttctgtat	51240
gagaaaaatc	tagctaatta	tttatttcag	tgtctctgga	atgcaagctc	tgtcctgagc	51300
cacttagaaa	acaatttggg	atgacaagca	tgtgtctcac	aatgctgctc	tggttgccag	51360
tgctgtgctg	ccagttgtca	tctttgaaca	aactgatgca	gtgctggttt	aactcttctt	51420
ctttttggag	taagaaactt	tggaggcctg	tgtccttcta	gaagtttgct	gagcaaatgg	51480
taaggaaaag	aaatagggtc	taaggcctga	ctatttcaga	gaatttcttg	atttattgga	51540
ctgtcaatga	atgaattgga	atacatagt	gtaggctgtc	ttttcttctc	agacactgca	51600
atttctccca	atctcttgac	ttttctagaa	gttttaatcc	aagtccttgt	tggttggttna	51660
gataaaagg	tattgttcta	ctagagactg	accttggcat	ggagatctca	tttggaactca	51720
cagatttcta	gtctagcgt	tgggtttgta	tccatacctc	gctactgcat	tcttagttcc	51780
ttctgctcct	tgttcctcat	gcccagtgct	ccaccctacc	cttgccccta	ctcctctaga	51840
ggccacagt	attcactgag	ccatttcata	agcacagcta	ggagagttca	tggtaccaa	51900
gtgccagcag	ggccgaattt	tcacctgtgt	gtcctccctt	ccatttttca	tcttctgccc	51960
cctcccagc	tttaacttta	atataactac	ttgggactat	tccagcatta	aataagggta	52020

actgctggat	gggtggctgg	gatacacaga	atgtagtatc	ccttgttcac	gagaagacct	52080
tcttgcccta	gcatggcaaa	cagtcctcca	aggaggcacc	tgtgacaccc	aacggagtag	52140
gggggcggtg	tgttcagggtg	cagggtggaac	aaggccagaa	gtgtgcatat	gtgctgacca	52200
tgggagcttg	tttgtcggtt	tcacagttga	tgccctgagc	ctgccatagc	agacttgttt	52260
ctccatggga	tgctgttttc	tttccagaga	cacagcgcta	gggttgctct	cattacctga	52320
gagccagggtg	tcggtagcat	tttcttggtg	tttactcaca	ctcatctaag	gcacgttggtg	52380
gttttccaga	ttaggaaact	gctttattga	tggtgctttt	tttttttttt	tttgagacag	52440
agtctcgctc	tgtcgccatg	ctggagtgtg	gtggcacaa	cctggctcac	tgcacctccg	52500
cctgccagggt	tcagcgatcc	tcctgcctca	gcctcccaag	tagctgggac	tacagggtgcc	52560
tgccaccatg	cccagctaat	ttttgtattt	ttagtagaga	cggggtttca	ccgtattggc	52620
taggatgggtc	tcgatttctt	gacctcgtga	tccgcctgcc	tcggcctccc	aaagtgtctg	52680
gattataggtc	ttgagccacc	acgcctggcc	gatggtgctt	tttatcattt	gaaggactca	52740
gttgataaac	ccactgaaaa	ttagtatgta	aggaagttca	gggaatagta	taagtcactc	52800
caggcttgag	gcaaaattta	caaagtctgc	tgactttgta	tgtaaagggga	ggcattttct	52860
tagaanaagn	agaggtaggt	ctctgggatt	ccagtatgcc	atttccatcc	tcagtgtttt	52920
tggccacctg	agagaggtct	attttcagaa	atgcattctt	cattcccaga	tgataacatc	52980
tatagaacta	aaatgattag	gaccataaca	cgtagctcct	agcctgctgt	cggaacacct	53040
cccaggtccc	tctttgtggg	tgaaccaga	ggctgggagc	tggtgactca	tgatccattg	53100
agaagcagtc	atgatgcaga	gctgtgtgtt	ggaggtctca	gctgagaggg	ctggattagc	53160
agtcctcatt	ggtgtatggc	tttgagcaa	taactgatgg	ctgtttcccc	tcctgcttta	53220
tctttcagtt	aatgaccagc	cacnggcgtc	cctgctgtga	gctctggccg	ctgccttcca	53280
gggctcccga	gccacacgct	gggggtgctg	gctgagggaa	catggcttgt	tggcctcagc	53340
tgaggttgct	gctgtggaag	aacctcactt	tcagaagaag	acaaacagta	agcttggggt	53400
tttcagcagc	ggggggttct	ctcatttttt	ctttgtggtt	ttgagttggg	gattggagga	53460
gggagggagg	gaaggaagct	gtgttggttt	tcacacaggg	attgatggaa	tctggctctt	53520
atggacacag	ractgtgtgg	tccgatatg	gcatgtggct	tatcatagag	ggcagatttg	53580
cagccaggta	gaaatagtag	ctttggtttg	tgctactgcc	caggcatgag	ttctgatccc	53640
taggacctgg	ctccgaatcg	cccctgagca	ccccactttt	tccttttgct	gcagccctgg	53700
gagccacctg	gctctccaaa	agccccta	gggcccctgt	atttctggaa	gctgtgggtg	53760
aagttagtta	gtggccccac	tcttagagat	caatactggg	tatcttggtg	tcaatctgga	53820
ttctttcctt	caggcctgga	ggaatataat	aactgagact	tgttttattt	ctgcagaggg	53880
ttctaagcca	ttacttccc	agatgggcca	ataatgcttt	gagtaatctg	gagatcatct	53940
ttaatgcgca	ggtgaatgga	actcttccac	agagggatgt	gagggctgta	gagcagagtg	54000
aactncctga	aactcagacg	tcagctcttt	gtctctctat	ctctgaacac	ccttccttag	54060
agatcccatc	tctaggatgc	atttctctgt	agttagtttc	taagtctctt	gttctctgtt	54120
tgcctttatt	tttttttctt	ggattctaag	ccagtatccc	cacttggtctg	tcttaatgta	54180
gcttaacatg	tctgtaatca	aaatgatcat	ctttctgaga	ttcaaagggc	tataagggac	54240
tttgagagaga	atttcattca	gttttctcca	aactagaata	atgcttgcac	tgtctgtaaa	54300
agaacaaaag	tgtcaaagca	tccttttggt	cactaaattt	ccttttttat	tatagtgtta	54360
cttaaatatt	aggaagttaa	aagttagtat	aaacttcntt	ataggctggt	attatacaac	54420
tatatgaccc	atacatattt	acaaattaa	tgacgccaaa	attgcaaaat	caataccatt	54480
caaattaata	ccttaaatgt	ggtgaggcag	ctgttgttca	actgaaacca	aattataagt	54540
tgcatggcag	taaatgctat	catgctgac	attttgagtt	tggccagtct	atatntatca	54600
tgtgctaata	attgaattct	ccaccatttt	ttctacttgt	atgaccttaa	tttgatggca	54660
cctgttccat	cctcatgagt	ttgtacaat	tatactgggtg	ccaacacaa	cataaacaca	54720
aatataaact	tgggctttga	aatcttgtgc	cagaacttgg	cttttaaagta	agcattttaa	54780
aatccatat	gtgtttatta	gactttgttt	agatgactgt	tgaatgaaa	acaaagtgtt	54840
taaaatcctc	ttagagaact	taaatataat	ccctcagcaa	tatgtataca	gatcttcctt	54900
tgagaaaaac	tgatttgtgt	cagcctctca	tgttacaaat	ggggaacctg	aattctgagg	54960
tctctagtga	gagaacaggg	actggaatct	gtggantcct	atctgtttta	ataataattg	55020
taaagtataa	tagataatat	tatattaata	aaataaaagc	aaacacttag	aatgagcttc	55080
catgtgtgag	gcactaactg	attaggcatt	attaactaga	tttattcctt	ttaaggcccc	55140
gcgatgtact	gttattttcca	catgttgtag	ctggggaacg	tgtactcag	agaggttaag	55200
taacttgtct	gaggtccaca	ccactaacia	ggagcacagg	tagggttcaa	atccagataa	55260

gtggagcaat ctcggctcac tgcaactttt gcctcctggt ttcaagcgat tctncctgcc 58560
 tcagnccctcc tagtagctgg gattacagac gccaccacc atgccagct aantttttgt 58620
 attttttatag agatgcgggt tcaccatatt ggccaggctg gtctgaaact cctgaccttg 58680
 tgatcctcct acctcggctt cccaaagtgc tgggattaca ggcgtgacct acagtgcctc 58740
 ctgagaattt aattttattt tatgtgcaag aggatccct gaggtagtca ggccacattg 58800
 tctggtgact cttgggatag agggaaactg aatgacaaag gcccaagaaa gcaattgtaa 58860
 tcattacata tacatggacc attttatgct gttttcttct ttcatttaac attatttagt 58920
 gtgcgtgttc acatatttct aaatcatctt ctgatttaga ataatgattt ctgatgtgta 58980
 ggctgtgttt tatagttttg aaagtaatac tttgatattc attacttntc ttgattctca 59040
 cagcaattct gnagggtgat gcgttgcaat ttctgtttca cagatgaaga gagtattgtt 59100
 aataagttaa tggccgggca tgggtgctca cacctataat tccagcactt tgggagacca 59160
 aggtgggagg atcacttgag gccaggaatt tgagaccagc ctggtcaatg tgggtgacacc 59220
 catctctact aaaaatacaa aaattagcca ggcgtggtag cacttgccctg taatcccagc 59280
 tatttgggag ggtgaggcag gagaatttgc ttgaacctgg caggtggaag ttgcagttag 59340
 ccaagattgc accactgtac tctgcctgg gtgacagagc gagactctgt ctcaaaaaat 59400
 aaaaagtgtc taagaggagg gctgggatct tttggctcca aatctactgt gggatgatgc 59460
 ctttgacatt cctgatagct gtgcagtaat ccattaacac agtttttata agttcaaanc 59520
 cctgttgcca acatttagat tgttccatgt gtgctgttac aaataaatta ctataaagat 59580
 tctatacatt taatctttta ttatttttgt attatttctg taggccaaaa tctgaggaaac 59640
 aggattacta ggttgaaggg aaatggccct tgaagtgtct gatcagatgt ctttccagag 59700
 gatccaacca atttaaatag ccaccatcaa tgcagagac tttgtagtgc aggggaaggc 59760
 ggctgtgttt taaaaatcat ttcccctctc tagcattttt ctgatgtgat ccttaagatt 59820
 tcacttttagt tttcccaggt ctcatggca tgtatgctgt tagggatggg tctaaaatta 59880
 atttttcttc acattcatac catgtcatcc cagtgattat ttaataaata atcacttgat 59940
 taaatagtga ttccttttct agttattttt gggacattta ttaaacctg gatatgggtg 60000
 ctcatgcctg tattcccagc actttgggag gctgaggtgg ggggattgct tgagactagg 60060
 agttcaacac cagcctgggc agcatagcaa gactccatct ctataaaaaat aaggaaatta 60120
 gtcaggcatg tnggtactt gcctggagtc ccagctactt ggaaggctga ggtaggagaa 60180
 ttgcttgagt ccagggtggtc aaggctgcag tgagctatga ccatactact gtactccagc 60240
 ctgggcaaca gagtgaact ctgtctgaaa aaaaaaaaaa aaaaaaaaaa aaannnnnnn 60300
 aaaaaaaga tgtgtaggga gcaatttttg agttattcat ttggtcattt gatatgtagt 60360
 ttttagttttg gtgctgatag agcccagaat gtaccctgaa tttgatgaac attctgatat 60420
 atgggggagc tcattgtccc ccacttacct ttttgctctc cagaatatct tttgatattt 60480
 ttatctgttt tttcccatt gaatgttatt accttatcaa gctcaaaaaa gtaccctatc 60540
 gctattttta gttaggtgt gttaaactta taaattagct tgggaaattt ggatattaaa 60600
 tgaactcatg aagaagcaga gtttagctct ccttaattct catcttcctt tatttatcct 60660
 actacagttc tgtggttttc ttttatgtaa gaagcacatg ntttggctaa gttaatgcct 60720
 aggttttttt gtttatgtgt ccattctcac tgtggatagt attctctttt cccacatta 60780
 tattaattta actggttttc agagactaat agcaatgcta ttatttagga gaatttacct 60840
 tggttctgat taacttaccc atacttgcaa atcatttgca gcttttttagt taactttgtg 60900
 agttctctta gatttacgac catgccagaa acagaaagga tattttcatc tcttctttc 60960
 tgatgtttat tcttcttgtt tccttttttt atccccatt atattctcaa gaatctctca 61020
 atactaagaa atagcgactt catttttcag cgnccgagtg cattattttg gctaccatga 61080
 ttcagaagcc tcttgccata ggcccaattt tattctgcta gttttctctg ttctttgtac 61140
 atggcccttg cgtgccccta accttgaatt aacgtggcta aatctcaaga atttaagagc 61200
 accgtgactg tgtcctcagg ctaggagggg aaatgggttc acagagtgcac tggattgtgg 61260
 tctatgaact tcggcagcca gcagcaaaag tcaggcatga ataataagat ggacagtga 61320
 catctgtagt gtgggagatg ttggcataac tatgaatgat gattcaagag tggtttgatg 61380
 catattgaat aacatgatga taagtactag actctgtgct aagccttcta tgtgaaatac 61440
 atttaattct cataataact ctagagcagt ggttctcgac cggggccggg tatcccccta 61500
 cccaccccca cctcaccct tccaccaggg acataacatc tggagatatt tttggtgtgc 61560
 acaatcctgg gaatgtatgt gctgatattt agaggttag gtcagggatg ctgctgaact 61620
 tcgtagaatt cataggagag tctctcaca cacctatctg gccccaaatg tcagtagggg 61680
 cactatcaag aaaatctgct ctagcagtgc ctgctcatat tatccccatg ttgaaatagc 61740

tgctcattgt	ggaaaatttta	gaaaagacag	gaaagtatta	aaanacatca	ctttttttttt	65040
tttttttttt	tttttttttt	taagnagaca	gagtccttgct	ctgtcgccca	ggccggagtg	65100
cagtggcggtg	atctcagctc	acagcaacct	ccgcttccca	ggtttaagt	attctcctgc	65160
ctcagcctcc	caagtagctg	ggagtacagg	catgcaccac	cacgcccggc	taatttttga	65220
tttttagtag	agatgggggtt	tcaccatggt	ggccaggctg	gtctcaaact	cctgacctca	65280
ggtgatccgc	ctgccttggc	ctcgcaaagt	tctgggatta	taggcaggag	ccactgcgcc	65340
agccacacct	acgttcttat	catcctagta	catccactgt	cattatcttg	ctgtatttcc	65400
ttctgcccag	tctcactctg	atcatgcagt	ggcgtgatca	tgcagtgatc	tcggctcact	65460
gcaacctagg	ccttctgggt	tcgagtgatt	ctcctgcctt	agcctcctgg	gttcaagtga	65520
ttctcttgcc	ttggcctccc	aagtagctgg	gattacaggc	atacaccccc	atgcccatct	65580
aatttttga	tttttagtag	acacagcggt	tcactaaaat	tttgtatttt	tagtagagat	65640
ggggtttcac	catgttggtc	aggctgggtc	ccaactcctg	acctcaggtg	atccgcctgc	65700
cttggcctca	caaagtgatt	acaggcatga	gccactgcat	ccatcgccaa	aaagattttt	65760
taaaagagtt	taatgtagaa	ccatatcaaa	ggtcttttga	aataaaaaaac	agttttttta	65820
aaatatcaga	aataaaaacaa	caaataaata	aataaataaa	aacaccccaa	acaatctgaa	65880
gcacgagcac	ctagcagaaa	ggttcaatta	tgatctattc	atagagtggg	atatcaagta	65940
gacattacag	gacatgtttt	aagattatat	tttatgtcat	gggaaatgct	ctcccagtat	66000
gatgttaa	gaaaaaacag	aatacaaaaag	tatatatgct	gcatagtctc	aatattgtag	66060
agaaaaata	ttatttatgt	atgcatgaaa	aaagacaaaa	gatgttaaca	gagatccatt	66120
gttacttcag	tttactaggg	attgtctctg	ggaggtagga	ttaaggatga	ttatattttac	66180
cttttttaaac	ttttctgtat	tttttttattt	tcaaattttc	cataaaaaata	taaggacttg	66240
aagatcaaga	aaaaatttct	gcttttggtc	agtgcagtgg	ctcacgcctg	taatcccagc	66300
agtttgagg	ccctagggga	gaggatcact	tgaacccaag	agtttgacgt	tccagtga	66360
tatgatctcc	ggatcgtacc	gcctggacga	tggagcaaga	ccctgtctca	aaaaaaaaaa	66420
tctttgcttt	ttttttttgt	ttgtttttga	gacggagtct	ctctctgttg	ccccagctgg	66480
agtcacagtgg	cacaatctca	gctcaccgca	acctctgcct	cctgggttca	agcgattctc	66540
ttgcctcagc	ctcccaagta	cctgggattc	catgcaccca	ccactatgcc	cagctacttt	66600
tttgtatttt	cagtagagac	agggtttcac	catgttggtc	aggctgggtc	cgaattcctg	66660
acctcagctg	atccaccggc	cttggcctcc	caaagtgtg	ggattacagg	catgagccac	66720
tgtgcccagc	ccaatctttt	gcttttttta	aaaaaagaag	acaaaaagg	attttatacc	66780
agtattatct	tggctgtgtg	actctgaagc	cacagtgtga	agttataatt	actctgaaac	66840
acaaggccct	gtgactcttt	tgggctcttt	ggtgtttatc	ttgattacaa	cgttggaata	66900
tagaaatgaa	aggaatggga	gaggtgatag	acttcaggca	gtgtaactag	ttgtctgaac	66960
actactggct	caattatatt	gtgtctagt	atctccatct	tgtccgtctg	ctaattttatc	67020
gcctggtaac	tcactgaggc	agggttttcc	tttggagaaa	cctcattggt	ttaaccagtg	67080
tatcatgctt	gtttagaagt	tcaatgatct	ttttaactca	tcggagaaga	tgatgaccag	67140
acctggacag	atggggaagg	actttgcact	ctctctttac	agtcctgagt	gcacacaggt	67200
caatatggaa	ctatgtgtga	atcttccattg	tctttgagag	ccctcttctc	tgccccatag	67260
ggagcagctt	tgtgtgcaat	tagaggagca	agggttgtgt	gtatttagca	cagcagggtg	67320
gcctggctct	ctcctctcaa	catagtcacc	acataacctg	cactatgcta	aggctgggaa	67380
tgcagacaga	tgggtgcctg	ctttcagagt	gctcaatgtg	ctgagggaagc	cagcaacaga	67440
aacagatgat	ttcaggagct	ccaggaaaaat	gctacaggag	gagtgtgcct	gggttactgg	67500
agtagcacag	gaggagggtc	cttagctcag	gctgagattt	tagtaaaagg	aattatgcca	67560
cgatgaatcc	tgaagaatga	atagaagtga	accagataaa	gcacgatagg	aagcatcttc	67620
ccttacctaa	gggaagacac	agaggtatat	ggaatggtat	gttaaaagg	tgggactcca	67680
aacagttctg	ttaaagctta	gagagtgggtg	ggagagactg	gagaagttga	ttaattagta	67740
aatgaagttg	tctgtggatt	tcccagatcc	cagtggcatt	ggatatccat	attattttta	67800
aatttacagt	gttctatctt	atctccact	cagtgtcagc	tgctgctgga	agtggcctgg	67860
cctctattta	tcttctctgat	cctgatctct	gttcggctga	gctaccacc	ctatgaacaa	67920
catgaatgta	agtaactgtg	gatgttgctt	gagactcacc	aatggcagg	aaaatccagg	67980
caattaacgt	gggctaaatt	ggacttttcc	aaagatgctg	tctttgggaa	acatcacaca	68040
tgttttgat	cagnaaaacc	taggcttcta	atgtgttgat	aaggcatgaa	ctcaggagac	68100
tgttttcagt	cctagtgaat	ggtgataatt	gtaattataa	cagtagacaa	catctctttt	68160
acacatttta	aatcatgaaa	atagaataac	cttactgata	attttagaaa	gtggtgatta	68220

aaagcacatt	taagataatg	ccttaacacc	tagtcttttc	catatgcatg	atgtcttaat	68280
cacacattgc	aaatcatgga	acacagaatt	ttaagcagca	tttgtgtaga	acttctcagt	68340
tttactaata	ttatttttatt	ttattctcat	aacaaccttg	aatagaactc	agatcatctg	68400
tcaatcatgt	atthttgataa	cagccttttac	agtgaacata	gaaaatacag	tagtggctaa	68460
caacacaggc	tccagatgtc	aggttatctg	ggatgaatt	ctgggtgtcag	cattcactaa	68520
gcataatgacc	ttggacaagt	gatttaagtt	tcttttaaac	agagaatagt	aatacctacc	68580
tcatattatt	attgtcagtg	tatcatctta	caatcacagt	ctttctctta	gggctgggct	68640
cagtgggtgg	attgacactg	cagaaatggc	cagatctaaa	ggatcaacat	ttacgtagct	68700
gggaaatgta	gctgggactt	cagtttctact	gccctagtga	tttttcctac	cactaagcag	68760
ctcagtcctat	acccctacga	gaccacacag	cttatgagat	actgttcttc	caggaaagca	68820
gtggggccag	ggccaccttt	taattgtgtt	tcttggcctg	gtcccatctt	tctcacaata	68880
tatagcaaca	gttattttact	tgctgatttt	ctaatagcaca	tcacacatag	tcatattaaa	68940
cacacacaca	cacacacaca	cacacacaca	cccctcaaga	aacattttct	gagacgtgat	69000
ttcctgattt	catcaaaaaa	gaaaagagcg	ggccaggcac	agtgggaagt	caagggtggg	69060
ggatcacttg	aggtcaggag	tttgaaacca	gcctggccaa	cacgggtggaa	cctcgtctct	69120
actaaaaata	caaaaattag	ccaggcgtgg	tggcgccacac	ctgtaatccc	agctactggg	69180
gaggctgagg	caggagaatt	gcttcaacct	gcgaggetga	ggttgcagtg	agccgagatt	69240
gcgcatttgc	actccagcct	gggcaacaga	gtgagactct	gtctcaaaaa	aaaaaaaaaa	69300
aaaaaaagca	taaactgaaa	tttatatgca	atttatatgc	ctgtgagata	attctgtttt	69360
ctcttttggg	accccaaaga	gattttttttg	attgatgagc	aaatacattt	tagattttat	69420
ttaagcatta	tgccaagcac	cactgaagta	taagtttcaa	gggcaaactc	agttttttca	69480
tctactagac	gaatgatttt	ctggaatgat	tacaagcagg	caangatggg	gnntagtggg	69540
aatagcaaat	gtcttcggca	tcagacaagt	tggggtttgt	ttgtatcctg	cctctgccct	69600
tcaccgaggt	tgtgatcttg	ggcagattgt	tgagttttta	cctagattcc	tctgactcca	69660
gatcataaat	tttcagaaaa	gttctgaaat	tcttgatat	actgatggta	aatgagactt	69720
tctcttcat	ctatgcactt	ctttgtttgt	ttgttttgag	atgggtcttg	tctgttgccc	69780
agactggagt	gcagtagtgc	aatctccgct	cactacaatg	tctgcctccc	aggttccagt	69840
gagcctcctg	cctcagcctc	ccaaatagct	gagactacag	gcatgtgcca	ccacgtccgg	69900
ctaattttttg	tattttttagt	agagacaggg	ttttgccatg	ttgaccacac	tgggtctcgaa	69960
ctcctggcct	caggtgatcc	gcccgcctca	gcctcccaaa	gtgctgggat	tacaggcatg	70020
agccaccatg	ccgggccata	tccatgcact	tcttgcaacc	ttaccttctt	ttctcatcac	70080
cctccaggga	cctagttgga	agagcagagt	taaaagttaa	ggtgaaactt	ggagaggtgt	70140
cttgtcccta	ggaacaaagg	actggtttga	aattctctgt	aaatcttccc	cagttcaaac	70200
cagagttatc	aaggtcttaa	aaacttccct	gggtcctgag	agcccattat	attattttact	70260
tgtcttctctg	tacaccctct	gcctagtcct	gatcctactt	ttgtttgcaa	ataggatggg	70320
gcacaacgta	caaggaaggg	cctttgccac	ccctgctaag	ggataacctg	aaataccttc	70380
accatcactg	ccctgtgctg	cttttcacct	atgccagtct	gtctacagtg	ccagtgtctc	70440
ctggcattga	aaggggagaa	tcttttggtc	ctttgagtat	ttgggtgggt	tacataaatc	70500
tccctgaatg	aagagcagct	gacttaggca	aggggccttg	tttggttttc	cttgaactat	70560
taacaggaag	ataggggagat	taactgtgta	aatgttcaat	aggccagagt	ccctgcagag	70620
gggtggccaca	gtgatcagat	cttatcacat	ccttgctttg	ggtgttgctt	ctctggttgg	70680
agtatggata	gaaaagaaa	aaagacccta	tattgaaatg	caaagtgcag	caagtcctga	70740
ctttggatta	acttctcagc	ccattttgcat	gaaaataaaa	agatgaataa	aacaaggttc	70800
ccacttttga	gggaggtggg	agctgtgaga	tgggaaggagt	gttcctgctg	ggcaacagca	70860
gagtaagtgc	tggggtagat	tcactcccac	agtgcctgga	aaatcctcat	aggctcattt	70920
gttgagtctt	tgtctacac	caggcactct	gcaaaaacgc	tttgcttgca	aggtctcatg	70980
cgatgctcac	cacagctctg	tgaagttaat	tgtactttta	tcaccatttt	acagatgaga	71040
aaactgaggg	tatgggggtca	atgacttggc	taaagtcaat	gcttagcaag	ctgcagggac	71100
tggatgtgaa	ttccaattgg	tttgactcca	aagcctgtga	agctacttgt	tcttcaccac	71160
ctagagctgt	ggttcttgat	aactgtgaac	tcttttgggg	tcacaaatag	ccctgagaat	71220
atgatagaag	caggagctct	ggcctttctg	tccatacctg	aacaggctcct	tgggttaaga	71280
gcccctcgtc	cagggcctat	taatcttgat	cctcataagc	agcatccatg	tattacggcc	71340
gcaaaccaaa	ctgtgccaga	ccgaatccta	ggaccaagcc	caaatatgtc	ccatcatcct	71400
tttggtaaga	agctcattgt	aagaaagaaa	gaggagagca	agaggatgac	ctagtgcag	71460

```

gggcctcatt gttttaatta gtgacaaaac aacaataata acaacaaaac ccccgaaagct 71520
tcacagatga catcagaccc caagcctgtg tgtttttcag gtgcccttga ggagctttgt 71580
agctggcaga ggaggtgaaa ctgacaaatg tttggcagat ggaggagagt accagagggg 71640
tttgagatga gctaaattcc aatctaaccg cagggtgtgag gaagagggtt ggattgggac 71700
catggagatg ggggttctac tcccagtcac gccagctgac tttgcgagtg ttctttgtca 71760
gtcactttat cttattttat ttatttttat ttttttgaaa tggagtttcg ctcttgtcgc 71820
ccaggctgga gtgaaatggc gcgatcttgg ctactgcaa cctccccctc ctgagttcaa 71880
gcgattctcc tgcctcagcc tccagagtac ctgggattac aggcgcctgc caccaagccc 71940
atcgaatttt tgtatgctta gtagagacag ggtttcgcca tgttggccag ggtggctctg 72000
aactcctgac ctgaggtgat ccgccacct tggcctccca aagtgtctggg attacaggcg 72060
cgagccactg tgcccagccc acttcactct accgtagtta cctccttaga gtatgaaaaa 72120
ataggcttag ggcatcccca agtccccctc atgtctgaga gctgaggctg gctgtcaaag 72180
aggaactaag gatgccaggg actttctgct taggaccctc ctcatcactt ctccaacgct 72240
ggtatcatga accccattct acagatgatg tccactagat taagaatggc atgtgaggcc 72300
aagtttccac ctgagagtca gttttattca gaagagacag gtctctggga tgtggggaat 72360
gggacggaca gacttggcat gaagcattgt ataaatggag cctcaaaatc gcttcaggga 72420
attaatgttt ctccctgtgt ttttctactc ctcgatttca acaggccatt ttccaaataa 72480
agccatgccc tctgcaggaa cacttccttg ggttcagggg attatctgta atgccaacaa 72540
cccctgtttc cgttaccoga ctccctggga ggctcccga gttgttgga actttaacaa 72600
atccatgtaa gtatcagatc aggttttctt tccaaacttg tcagttaatc cttttccttc 72660
ctttcttgtc ctctggagaa ttttgaatgg ctggatttaa gtgaagtgtt ttttgtaaat 72720
gcttgtgtga tagagtctgc agaatgaggg aaggggagaat tttgtgatat ttgtgaacat 72780
tggggatatc atcacctcga gtatttatca tttctgtatg ttgtgaacat ttcaagtcct 72840
gtctgctagc tattttggaa tatactatat gttgttaatg atatcatgca gcagacgtgc 72900
aatcgaatgg gctggctcta ggagctagag ggtaggggct ggcacaaaga tgcagtctgg 72960
aagggtcctt gcccataaga agctgacagc caaggctagg ggagtctgtt cttctctgca 73020
tcaggtcacc tctctcacct ctgtcactgc cccatcagac tacaatgtct gcaggctctt 73080
ctccccagag tgtgagctcc ctgagcaaag caggatgctg ccccttccct ttgtattcct 73140
ggctcctggc ttcagtgcct ggacataagt atgggcataa taagtgtccc ccaaatagaga 73200
cattgaggat tcttcaaagt cacaggaccg tgatgtgagt taggacggag taaggacgat 73260
gggatgtggc tcaggacaat cctgaggaag ctgcagctgc ggcacgcagg gccacactgt 73320
catgttcatg gaccctagac tggctttgta gcctccatgg gcccttcca tacacaaata 73380
ttaaaaatta tatttcatga ctgncattgg tataaagatg aatataatcc agaccagatt 73440
catgattatt catacatttt tagtgtatta acttttaatt ctgcttttaa aataaattaa 73500
aacattctaa tatgccctta agagtatccc agggccaggc cactgagcct actgtggttc 73560
atggataagn ttggccctgg gggcatgtgt gtgcatgcat gtgtgtgcac atgcatgatg 73620
agccgggcct tgaagggtgg taagatttgg gtgtgtagac caatggagaa aggcatttgg 73680
ggcagtgatg atgggtgggg gagggaaacat ggtgatgaat ggagctgggt gtggggagcc 73740
atgggagtgg gttagggcca gcctgtggag gacctgggag ccaggctgag ttctatgcat 73800
ttggcagtca cttctgtaaa gcagcagagg cagtggcct agctaaagcc tttcgccttt 73860
tcttgcaccc ttacagtggt ggctgcctg ttctcagatg ctcgagggtt tcttttatac 73920
agccagaaaag acaccagcat gaaggacatg cgcaaagtct tgagaacatt acagcagatc 73980
aagaaatcca gctcaagtaa gtaaaaacct tctctgcatc cgtttataat tggaaattga 74040
cctgcaccag ggaaagagag tagcccaggt gtctggggct tgttccattt agatcttccc 74100
caaggggttt ttctccttgg tggctggcct gtggggcccc tctccaggag gcattgggtg 74160
agaaactagg ggagctgggt gccacagaca gtgatgtact aatcttctct gggaagacag 74220
aagaaaagtc cccagggaag aatactacag acttggcctt agggacagct aggggtgcag 74280
attgctgcca actgcatttt ttctgaagtt ggccatatgg ttgcagtga tggatttata 74340
gacagagtat ttctgtgcat ataagagcaa ttacagttgt aagttgatat ggataagtga 74400
aagttaagca cttcttttcta aaaagagaat gcaattcatt ttcccctaatt catttcaatt 74460
agtctgantg ggcatttgaa cttgttgtct ttaaaaagtg aaatctttac ctctgatctg 74520
gtaagtatcc aggcattttc ttgtgtgcca cccaggaggt atctggggag tgggcatttt 74580
ctgactgagg cattggctgc catagcatca gagcagcctt ccaggcagtg gcctggcaag 74640
gggacagagg ctggtgggag cagctggctg agtgcagcca gtaatggcat gtgcatggct 74700

```

tgtagagaat	gtagaagcaa	taatgaagcc	gataaaaagct	ggctctgcatt	ttattattat	74760
catgcgccgg	tgggttctaaa	caatgtcagt	gataaattac	tcctcccat	catggacca	74820
tggctgccac	tgctccaggg	aagtgccttt	tattccgttt	gggtgttagg	gagggatgga	74880
gttggctggc	ctttgctgaa	aggcctacca	gtttgttttc	tatttgga	aagaagaaat	74940
gataaagtty	tagagtttaa	accagactca	gatttgagtt	tttttttttt	tttttttttt	75000
tttttttaag	gctacagaac	tgtgctttcc	ttgggcagta	aaagaggcaa	tgggcaatgt	75060
gggacctgat	ngacaanagg	gaannngnnn	ncnnannnn	nnanngctgt	cttaggggtg	75120
gcatggagga	ggtgctgctt	cacagcagag	agaggtatgg	ctgtgcttgg	agtgtccact	75180
tagacaactc	ctggctgtgc	agccaggcca	tcgagatgct	gtttccttga	cctgcaggtc	75240
ctggtcttgc	acatggatgt	ttcttctggt	gcaggagaca	gaaaggtagc	aacaaccct	75300
gatcaaagcc	tcagtccttc	cttatttact	ggagagcccc	tgctgattga	ccagaggcac	75360
agctggggat	atttctctta	cctctgtagc	aagagacagc	gtggtgcaga	ggaaagtgt	75420
agcatacatt	accctgtggc	tgcagactt	tgtgaatagg	ttagttagca	ccctttcagc	75480
cacttcttct	tacctgttaa	tgagataaaa	catgtaattg	cttaaaaaca	gtatttggca	75540
cataggaagc	acttagtgaa	tatgaattat	gatttttttt	ggagtgggtga	catctcaacc	75600
aagccattta	accctcagc	cttnactttc	ctcaactata	aaatagcagc	taacttgaaa	75660
tgtaaactat	aaaaccta	gtagtatctg	gcacatagta	gattcccaat	aaatgagagc	75720
cagtattctt	tctaagacag	tgatgcattt	ctgagcacct	ggccttgctc	ttctgccttg	75780
caatttatgc	agcagttgaa	atagactggc	tgatgggggt	aagtgtgcaa	gcagactttc	75840
tgatcttagt	ggaggagact	gccttaaaac	aacactaatt	tcctttttct	tttcttttct	75900
tttcttttaa	gacagagcct	cgctctgtca	ccaggctgga	gtgcagtggc	gcagtcttgg	75960
ctcactgcag	cctctgcctc	ctgggttcaa	gcgattctcc	tgccctcagc	tcctgagtag	76020
ctgggactat	aggcatgctc	caccatgccc	aactaatttt	tgtattttta	gtagagatga	76080
gatttcacca	tgttgccag	gatggtctcg	atctccta	ctcgtgatct	gcccgcctag	76140
gccttccaaa	gtgctgggat	tacaggcatg	agccaccatg	cctggccttc	tttgagaagc	76200
tggagacatg	agttaagtgg	tgaagaagcc	aaatctgtat	ctaaaaacc	tacagtagtg	76260
tgcagagctc	tgaggagaga	aggtccctta	gattttgagt	gtattattat	gtcagtgtct	76320
gttttacatc	tctctgttca	cgcagtatgt	ccccttttct	gccttgagc	tgtttcttaa	76380
attctttctt	tctttgcttg	tcttgagca	caaaacaggc	ttcagtatag	ggggaaatgc	76440
acagaaacac	tgccttttcc	tacaggaaat	cagtaacttt	ttactgattt	tgtttttatt	76500
tacttatttt	atttgtttaa	atttattttt	agtttttttt	tttttttagag	acagggtctc	76560
tttctgttac	ccaggctaga	gtgcagtggg	gcccttagag	ctcactgagc	tcactgcagc	76620
ctcgaattcc	tgggctcaag	tgatcctcct	gccttagcct	cccgaagggc	tgggatgaca	76680
ggcatgagcc	actgcacctg	gccaactttt	tgctgattgc	gaatagcact	cttgtcaatt	76740
tcggagagaa	gctgagactg	gcatatgtca	gtatggatcc	ccacttagag	acctgtgttt	76800
atctgcactg	acaccccatc	acagcatgat	gagcttggcc	ctcctgtgct	gtctctccca	76860
gggctgggag	gatccttgaa	gctgatctgg	tttggagctt	tgtcctcatt	cacctccttt	76920
accacacacc	caccttccca	gggaggggat	ctaccactca	ctaagtagcc	cattctgggt	76980
gttgacagct	ctaattgtta	gaaaatattc	accacctgt	tatgctttct	agagaacaag	77040
tctaattctg	ttttccttga	aatagtcgaa	gacagctctc	atgtttttct	ttccctgttt	77100
tcccaaagtc	catgattttt	taggcaaaat	ggcctccttt	cctcttatgg	aatgtttttc	77160
ctccccattt	ctgcctctcc	tctggttgtg	tttcagtatg	tctgtgtgct	tcttgaagtt	77220
tactggaaat	tatgaaagta	ttctggcaca	gaggaggaag	ggatttttgc	ctcccttgtt	77280
ctgagtgcga	catttccgtt	aatgcagctc	gagattgtat	taggcatttt	ggcattcacg	77340
tcaccttggt	gactcatatt	ccatgtgcac	tcaaacaaaa	ttgtgattat	tttaaataag	77400
cagaattgca	agttacgtgt	tctccatttc	tttgttgtat	tgttggcttt	ttgaactaaa	77460
gggaaaaaatg	tcttttttct	gttttacatg	tttagattcc	ctatgctatc	ctatcctccc	77520
aaaaccattt	tagattctga	ttttgccatg	tattatatct	gatactcctt	tctcgtcatc	77580
tagagatgtg	ataaacaact	ctctttggcc	tcattccagt	catcgataac	tgtgggacaa	77640
agacttgaag	cttggatcag	tcagtgagg	actaaccacc	cctgtagacc	cttttttctt	77700
caactataaa	atagcagcta	acttgaaatg	taaactataa	aacctaatgt	agtatctggc	77760
acatagtaga	ttcccaataa	atgagagcca	gtattcttct	aagacagtga	tgcatttctg	77820
agcacctggc	cttgttcttc	tgccttgcaa	tttatgcagc	agttgaaata	gactggctga	77880
tgggggtaag	ttgtcaagca	gactttctga	tcttagtgga	ggagactgcc	ttaaaacacc	77940

actaatctcc	tttttctttt	cttttctttc	tttttttttt	ttaagacaga	gcctcgctct	78000
gtcnaccagg	ctggagtgca	gtggcgcaat	cttggtcac	tgcagcctct	gcctcctggg	78060
ttcaagtgat	tcttgattct	gtagacacta	ccactcaggc	ctatattgta	atcagtgtctg	78120
ggccactggg	ctcctgcttc	tgtgatccag	ttgggaagtt	tatcttgttc	ttcccttcag	78180
cttgcatctg	ctaaattcgc	tggactatac	acaggtgatt	tgtagatatg	gggatctcta	78240
ctcaaatact	ctcatgattt	ccttggttag	agcatcattt	tatttccact	tattggaaga	78300
gaccttagag	accagttagt	tcatattatg	ataaattagt	tgattctgtc	attcaaccct	78360
tatatattga	gcgtctccta	tatggaatc	actgttctaa	gtgccgagac	acagaggtgt	78420
ccaaaacaaa	tatggccctt	ctccatatgg	aatttctatt	ctagagaaga	atctgaccca	78480
gaagggggaa	gtgactgtcc	caagtctaca	caaccacaga	agggatattc	tgggaataaa	78540
tcacggctaa	acccccctgc	tgtccaggc	agttctcctt	ctacagtgtc	ctattgtgct	78600
gttttaataa	tctttcaact	gggaagaact	cccatctcag	gaattaagcc	gtggacaaat	78660
cttttaatta	tccttgaaat	cactctaata	agaaatccaa	ggaggaagtc	ttacaggggtg	78720
cctcaccac	ttttctcatc	actggaactt	tttagacatt	ttattatttt	cttcctaaac	78780
cagagtacag	gcacacaagt	tgagtggtgt	ggtggctaaa	ttaatattatg	tttgcaaggc	78840
agtgtgagaa	gcattcattc	atcttaaata	cctatgggtga	ctgcaactca	gatgtaaaaa	78900
ttggataaat	cctcagaaac	cctagggaaa	gtgacatgtc	tgtattttgt	ctctgtgaga	78960
tacagactgg	cagagataag	tgtttctctg	ggtgagtttt	gtgtggtatc	tgggatgatt	79020
ttaggcagta	cgtggatgag	aacttttaat	tttaaccac	atccaattgc	aatttcatgg	79080
aaattattgc	ttaggaggat	gttcaacagg	aaaaatataa	ttaaagttaa	ttcaaaagaa	79140
acattttctg	tgaatatggt	aaaacttggt	agagtagttt	gtaaatgatt	gaagattgga	79200
aaacattggt	ataagagtga	gtgtgggggt	ttgtattaa	attcattttg	ggaagaaatc	79260
catgctgcat	ccctcatgaa	gtgtgaactt	tgggcatgtg	ttgattcttt	ctggccaga	79320
gtttacctga	agattagctg	ccctgagggt	cactgagcat	taaattagat	gatgtctgtg	79380
gatgactgat	agtgaagctc	atagcccaca	gttgacacat	aataaattcg	agttgtcttg	79440
cttccccttc	tgttcctggc	tgactgtttg	gcctttgcc	cttggtcggc	ctctctgggc	79500
cttaagtttc	tttgccctga	aattggaact	ttctttggtg	aaacaaccag	aaaatgcttc	79560
agcccagaaa	cttggtcagt	acttggtggt	gggatcacct	gggactacc	aaggatgtgg	79620
gctgtctgct	agactatagc	cccttgaggg	caaggtgggc	gccttgctca	tggttccatc	79680
ctaagcccca	gcacagtagt	gggtgcatgg	tgagccatta	gtgaatcttt	gtggaatgaa	79740
ggtgggagaa	aataaaatac	ctgtacttca	cagggtattg	tgaggggtcaa	gtaaaagtgc	79800
tttaaaaaat	tgtattatat	agtttattcc	cttggttag	cccagggtcaa	cagagcctac	79860
gaataataat	gatgacagaa	gttcttcaaa	aagtctggc	cttctttctt	tcacaaaatt	79920
gccccccaga	gctttctgga	agggcagcca	tgaaccaga	ggcctaaagt	agatttactg	79980
ggaagctaaa	aatattttact	ttatttttca	tagctccttt	caaggtcctc	tctgggggtc	80040
ttagcaatat	gtttacacag	tggatgtttt	ttgtaagggt	tgcaaaaagta	agatttttta	80100
aaaatactat	cttggttttaa	aaagagagcc	ccctaccaac	ttgtgtcagc	ctcaggcccc	80160
cacctgcac	tgtctctgcc	agggcatggt	ggggcaagaa	gcactgctcc	ccttccaaag	80220
cttccttcct	tgccttgagg	tcactctcac	tccccactcc	aagccacctg	ccatcgctgt	80280
gccccctctc	tggatgaatct	ggcattctta	ggtgggtga	gaagcagact	ggccaagct	80340
aaggcctttc	tgatgggggt	gttgctgctg	agaatcatga	ctgggtggga	gaaggaggtg	80400
accctttngc	tgtcttattt	ttactgtgta	tttccctttc	agctacttaa	aatgtattgc	80460
ttagtatac	ctaattgggt	cattagcctg	cttctactg	aacatttccg	ctcaggcatc	80520
cacttggtcc	caaggcctgc	tcctctccca	tattctgaaa	tctggactac	agactctcat	80580
tcaactccag	gttgactgtg	ggacacagtc	ccctcttgag	caggtagctc	cttgcatggt	80640
ttgggacgct	ctacttggtc	catagttggg	aagtgcata	gctggagctg	aagcctcttg	80700
ccttcccggg	ataggcgctc	ctcacatccc	ctctgagaag	ttccccagct	tccctctggt	80760
ccccggttcc	acacttagcg	aggctcttgt	ccactgctac	atcccccata	gccagtccct	80820
cagccttgcc	attgcttatg	ctgggtctgga	aacaattcct	agacttggtg	ggcatctggg	80880
gaagttctcc	atcttttttt	agctggcatg	acccaagtgg	tgtgggcagg	gctgtggatt	80940
ctatggtgtg	gctggaagcc	aggtagcctc	tctctactgt	acatggaact	cagcaacttc	81000
tgagtcaagc	aagatcttag	ctctgcaggt	gtcttgccct	gtccaaagt	atggccacac	81060
cagtaccttt	taactctaga	agcccagtaa	gtgttttgtg	ggaccgcaa	gatcattttc	81120
tagacctgct	gaatatgcct	agaacgggta	gggatggctt	tcacgctgtt	cctagggctg	81180

acaagtcaca cgtttctggg ggtacatata caccgcggtc cctgtgaatg gcactctcca 81240
 tgagaactgt gatgatttga gttgaatagt gcacagccta catgggttctc tgccatggcc 81300
 tggagttcct tatcttgctt tctccagtga ggactagggc tgcaactggc ctactttggc 81360
 tcctgacttg ggggattctg aaataccttt tttttttaag gttgtggagc tctctgaagc 81420
 ttataaggat tttgccagga aaagataaga aatatctttg ggcattttgt cactgtgctg 81480
 gagatgaacc ctttgaggga catatcacct tgttgaggtc aaggggcgga aagggacagg 81540
 actggcagag agatccgggg cagcagcctg ccatcccgac tgagtatgga gtttctctct 81600
 ccttcagct gcacttttgt gtggagtcag tggctcagct gccacttccc ttatgttcat 81660
 ggcatgaatc tggcttggtt ggcccttctt tttttttttt ttttttgaga tggagtctca 81720
 ctctgttgcc caggatggag tgcattggag gtatgtggtg gatctcggct cactgcgacc 81780
 accgcctcct ggattcaagc gatttctctg cctcagcctc ctgagaagct gggattacag 81840
 gcgcagtgca caacaccctg ctaatttttt aatttttatt tattttattt ttttcaaggc 81900
 agaagaattt ttcttagtac agaacaaaat ggaatctcct atgtctactt ctttctacac 81960
 agacacagca acaatctgat ttctctatct tttcccaca tttcccctt ttctattcga 82020
 caaaactgcc attgtcatca tggcccatc tcaatgagct gctgggtaca cctcccagat 82080
 ggggcggcgg ccgggtagag gggctcctca cttcccagaa ggggcggcgg ggcagaggcg 82140
 cccccacct ccgggacggg gcggctggcc gggcgggggc tgccccccac ctccctcccg 82200
 gacgggggtg ctggccgggc gggggctggc cccacacctc ctcccgagc gggcggtctg 82260
 ctgccctgct aatnttttgt attttcagta gagatggggg tttaccatgt tggtcaggct 82320
 ggtctcgaac tcnnnntng acctgttgat ccacntggn cctcangcnn ctcccaaagt 82380
 gttgagatnt acaggcatga gccactgcgc ccggtctgt tntttttgtt tgtttgtttg 82440
 tttttgtttt tgttttgaga cggagtctta ctctgttgcc caggctggag tgagtggagc 82500
 ggcacgatct tgggtcactg caacctctgc ctcccgnggt tcnaaacaat tcttcctgcn 82560
 ctncagcctc ccgaatagct gggattacan ggcacttacc accnaggcct ggctaatttt 82620
 tgtatttttt agtagnnang acgggggttt gncatgttg gccatttgaa ctctgacca 82680
 tctttngaac tcctgaccat gtttgcnat gttgaactct tgacttcagn gtgcgttggc 82740
 cntcccaag tgcgtggatt acagggtgta gccaccatgc ccagcctggt tggcctttct 82800
 gatatgngnc ntcntgacta atctttttgg aaattnagnn ntccccaggg ttatactgga 82860
 ttttacttag ggaaaagggt catgcnctct ctggctgtca gnatttactg atagtactaa 82920
 ggnactcagn tgggggtgnnn gnacctttga ttcnntnggt ttgatttttg aaaatncaaa 82980
 aagacgtgag ctccaggggag cagggtggct ttggtganca tggcaagata gttggctgtg 83040
 gcnagggagt tgaggggaagt gggtagaaaa ttaacatctt gtanaatatt nncncctggg 83100
 aaatatacct tcgtgttaag agaacagact tggcagncn nannngnnan nntnnnggcc 83160
 taggttcaaa tcttggttg atgcttatca gctgtgtaac cttggataat tccatacatc 83220
 tctgtgcctc agtttctca aatggaataa caatagtacc tccctcagga ctattgtggc 83280
 aaattaatgg acgaataagg ggaagcactt agtacagtgc ctggcccagc ataggtacca 83340
 ggcttgttct taagctcact gcattttttac aatcatcata aaatgcaggg gatacacaca 83400
 tgaaggagcc gaagttcaga gaggccaagt aacttgctta agaaaagcaca gctggcaagg 83460
 ggcagtaata ggaccagaat tccagtttct cgtgctctgt tgttgttata tcctaaagag 83520
 agcagctctg agtagccaga agcttcccta aagtacagag acatggggca tgggctggct 83580
 gggatgagaa aggagacaag agggcttctg aaagaaatgc cagattcact ccacttctg 83640
 gcttcaggca ccgatggaat gtttcccaag gccatctag aaagaacatc ctgtgactca 83700
 cageccattc tcattttctg tctgaacccc ctacccatt caggcagctg ctaaaagtga 83760
 gggtacagcc tcagactata ttttctgtcc ttgtggaacc ccagtgtgtc atcttgttgg 83820
 gagatctggt gatacatgtg tcaacattat gtcataaaaa tggaaattct ttgaaatctt 83880
 taggtgattg caattcacgt tctgtatgta tgcacttgtc aaaagttttg atttgaggcc 83940
 ttagaatttt atatttgga acctttccac taccatgagt tttcccagac ctgtcaaagc 84000
 caggctgcat ctcaaaaacc aggtctttga tttccatcca gggcaagggc ctgggcccag 84060
 ctgggctgta agcagggtgg ggtggggagc aacgctgcac tgcaatgttg aaatattact 84120
 tgaactaaat caaatcaaag atcagcttta ctcaacaag aatagaaaac acaattgcat 84180
 tcgattacag aatagtgtgt atccccacaa atatcagact gcctttaaaa agttttgaat 84240
 tgtaaacatc aagaacagtg ttgcgtgtct cctgcttttc cagcataagg tttatttatt 84300
 ctgtgggtgg caaagagcaa tttgggagtc cagtttgttt ctcatgaaa gctttcccat 84360
 ttctggctct cttgtcactg ttgcattgag gcacaaaag gcaatctcag tgcgacacta 84420

ttcaacagac	taagttgcac	cggataatga	taccatttta	catttttcat	atattattac	84480
attgaaggct	tcaaacagca	ctagcagggg	caaattggta	ttattatctc	catttcattg	84540
atgaggaaac	cgaggctcga	aagggtaatg	tctgttgccc	aagattaaag	agtaaagttg	84600
tacgttgaat	cggggtctga	ctcctaggct	tagcattttc	tccccacact	atgctgccat	84660
gttgcttatt	ccaacattag	gaagcatagg	tgccatcccc	agcttttgag	gccaatatca	84720
cgatgaagca	tttttaaaac	atctcattaa	attgctgata	tagtggaag	aaccaaagct	84780
ttgcagtcaa	agctgttttg	gttcaaattt	accacttggt	tgttctatga	ccctgagcaa	84840
gatattctcc	agatctgttt	cctcatttga	aaaatgggaa	taataatacg	tttctttata	84900
ggctgttctt	aaagattctg	gaaaataatg	ctaatagtgt	gcctaattgt	tggtaaatat	84960
gagtcacttt	tctgtgcccc	caaagcacta	ctatgtccct	taataaattt	tgttaatttt	85020
aaaagttaga	aaaaaattaa	actattttata	cattgtgtat	gttaattcct	ccctagacca	85080
gccttaggaa	gaatctcatc	cccaacttgt	aaactcatct	ttttccgttc	tttttggtgc	85140
tggacttctc	agggccctgc	aggctgattc	tagtcccatg	ttgtgtgggtg	tttgaagtg	85200
ctggctccct	ttttcagtga	gagaccagct	catccttggg	aactgaatgc	ctcaaactct	85260
cttttctttt	tctctcttcc	ccttctgttt	gatgtagtct	ttcctgtttc	tggactctgt	85320
ttcttcatac	ttcctatctc	ttaccttctt	ttcactcctt	ttgtcttcca	gctgtcctct	85380
ctcatttttt	tgcctctctg	gtcttcaggt	agagttttca	tctcagctat	cttcttgctt	85440
tttctgatgt	tggttctttg	tttcttcttc	tcattctgtt	caggtcctaaa	attcatttgg	85500
gtcaatgtta	tgtcttagtg	ttatcttcca	tttcttctg	agcttcaaag	ccaggctgac	85560
tgtgcccttc	cacgccttgg	ccagtgtagc	caggacatcc	tttccctctg	gggctgcact	85620
ggctcttttg	gggaattgtt	accattcagg	gtcttcaacc	ctcattctag	ggacttccag	85680
taacttctcc	caccttccct	cttctcagta	agacatgggt	attgtcttat	tctgtttctg	85740
ctgctagaag	aaaattctcg	agactgagta	atttatacac	aatagaaatt	tacttctcac	85800
agtctctggag	gctgggaaat	ccaaaatcaa	ggtgttggca	ggtttggtgt	ctggtgaggg	85860
ctgctctctg	cttctaagat	ggtgccttgt	tgtgcacatc	ttaggagggg	acaaacgcca	85920
tttctcctca	tggcagaagg	gactgaagcc	tctccctcaa	gcccttttat	aggggtcctc	85980
attgtcaggg	ctctaagccc	aagccaagcc	atcgcacccc	ctgtgacttg	cacatatacg	86040
cccagatggc	ctgaagtaac	tgaagaatca	caaaagaagt	gaaaaggccc	tgcctcgcct	86100
taactgatga	cgttccacca	ttgtgatttg	ttcctgcccc	accttaactg	agtgattaac	86160
cctgtgaatt	tccttctcct	ggctcagaag	ctccccact	gagcaccttg	tgacccctcg	86220
ccctgcccc	ccagagaaca	accccttttg	actgtaattt	tccattacct	tcccaaactc	86280
tataaaacgg	ccccaccctt	atctcccttt	gctgactctc	ttttcggact	cagcccacct	86340
gcagccaggt	gaaaaaaaaca	gctttattgc	tcacacaaag	cctgtttggg	ggtctcttca	86400
cacagacgca	catgaaactc	atggtatctg	tgagggggccc	taatgctatt	ggtgagggca	86460
ctgctcacat	aacttaatca	ccccttaaag	gccctgcctc	ttaatactgt	cacattggtg	86520
atgaagtttc	aatgtatgaa	ctttgagggg	ggacacgttc	aatcatagc	aggtgtgcta	86580
ttgccttact	agcaaagtaa	tctgggggaa	gatgagtaat	gtctcgttcc	acccttgatt	86640
ccagtgctga	ctgattctgg	agtggcagag	aggggtggag	gtcaccgct	ctgctcacc	86700
gaccctctgg	ccatctcctc	ttagaatgca	agggcagggg	ttttgttaca	cagcgcctct	86760
tttgttatag	gctagctcct	gccttcaagg	agcactgaga	aaaatattat	ccctttgaac	86820
cacaacacta	gtatttttggg	tactgtagcc	attcagaagt	ttgttgaact	aatacaagtt	86880
tatattattt	gtaaaatact	agaaggaatg	tgtggttcc	aaagttatgg	tattcccatt	86940
ggttgaaaga	gaagancctc	ggcttcctaa	tgccttggaa	ggttcagggg	gcataagcct	87000
aaatatcctg	gtttctcttg	ggaattcttc	acagcttgtc	atctatgact	ttatccacat	87060
ttattttgaa	catgtttaat	tttcagcttg	tattttctca	aggagtatgt	gtgagtcagc	87120
attcagagta	cttaacacac	ttagtgtcat	tagtgataac	ttaaccaca	cgtataggaa	87180
gctagacctc	catttacaga	aaaatggagg	ctcagagagc	ctaactgact	ttcccaaggc	87240
cacataactg	gaaaattctg	gaactgagtt	ttttcaggta	tgtcttggtg	tcctcgggtg	87300
agccttctct	gtagagaagg	aagtgtctcc	tgggttggaa	agcccaggcc	gggatgaaca	87360
gcatgggtgt	ccttaggctg	tgtgtaccac	acacccttgg	cctactctcc	cctgctctaa	87420
ggagtcacct	gaaggagcag	ctgcatcacc	ccgcctcacc	ccttctcctc	agcacccaac	87480
aaccaggtct	tcccagagag	tgctcagcta	ttggtgaagc	aggcatcgtc	tattcttcaa	87540
gaagacagca	gagcatcagg	aaaacaaaac	acaaaaaaa	taccataaag	acactggagt	87600
gtatttgta	ccagcccttg	tataacagag	aagtatttca	aggtcttatt	tgtgaaggctg	87660

gacaagctca	ggctgattaa	aaccacgagt	tagaatttgt	ctcccgtcc	caggccctgc	87720
tcagttgaca	gctcctagca	ggctgggagg	ccccaccca	ccccacccc	aggggtccat	87780
tgaagaagac	catctgggcc	tgggatgctg	atgtgatttt	ccctacttct	cctttttttt	87840
catttgcctc	ctccacccac	ttctaaagca	gctgattgct	ctggttcctt	ttcatgttga	87900
gcaaaagaga	agacagatct	tgaaacctga	gctggggact	ggagggcact	gaccgagagt	87960
aaccccgctc	ggctcagctc	agctcagcaa	agggctttcc	tcctccgcct	tcctgagctg	88020
ctctccctcc	tctcctttcc	tccttctttc	ccagcctccc	ccaagctcct	tcacccctct	88080
cttccctctc	ctccctcctc	ccaccttcct	tgtcctctat	gctcagagaa	ctgaagggaac	88140
actgaatggc	ctggaaagca	ggacctcgcc	ccccacccgc	gtcatacccc	tccgcctgca	88200
cgtgcaaatg	tgttcctcag	cttgggtcca	gaggcctgga	cctgggtccc	agagggtcca	88260
gctggttccc	cggcttggtc	ccagaggcct	ggatgtgtgg	caagaagctg	aggggttctgc	88320
ttcgttttct	ccagcccaca	aacacacatg	tcacctgttt	tttgttcttt	ttcctttcct	88380
ttgccccact	ttacagaagc	tcttggctaa	ttggtaggtc	actgtttgct	cctccaggca	88440
aaaccaagga	ggctgagctc	agcaggtaa	gtattttggg	ctgtattgaa	gagactgtgg	88500
tgggagagcc	ctcttgacct	tctgctggat	agtgactttc	tcttctgctg	gcagcagtag	88560
cccagctccg	catcctgtga	cagaggatcat	accttccagt	tacccgagcg	aacacttgga	88620
gcaagtcagc	tttttccctc	gaaacagcac	gttcacgttc	ttgatctttg	gtttgcacaa	88680
caacctgtga	atgttattag	ccccgttttt	tttgtttggt	tgttttttga	gacagagtgt	88740
cgctcttggt	gcccaggctg	gattgcaatg	gcgtgatctc	ggctcactgc	aacctctgcc	88800
tccaagcaat	tctcctgcct	cagcttctctg	agtagctggg	attacaggca	tgcgccacca	88860
cacccggcta	atthttgtatt	tttagtagag	acgggttttc	ttcatgttgg	tcagggtggg	88920
ctcgaacttc	cgacctcagt	tgatccaccc	gtcctggcct	cccaaagtgc	tgtattagcc	88980
ccattttata	tccaagga	ctgaggctta	gggagattta	aatagggggg	aaagtatagc	89040
ttcatgatta	cacagtcagg	aaatggcaga	gctgactctt	gaattcagac	cctgtgatgc	89100
tgtatccagt	gctgattttg	ttttgccttg	tgttcctttc	agtgtcaggt	aagaagacct	89160
cagagatcat	cttctccaac	ccacacattt	tactgattat	aaaatccagg	tccgcattaa	89220
ctcattggcc	aaagttaagt	ggctaaatga	ctctcctggt	caggaagtat	tccctgtctt	89280
gtgtggaatg	attgggtctt	cgtccaggct	atcttccacc	tatatccatc	cttgcatgta	89340
ctgggtggata	tgatcattcc	tggtgattcc	ctgtaaaata	ctagatttat	ttattctttc	89400
ataaagttat	gagtagacaa	atcagtcagc	ttcaaccttt	cctttacctt	atagatagta	89460
tcagcagaaa	gaactcaata	attctgtctt	aggaactcag	ccaggcta	atgtattata	89520
atacacacat	acatatgcac	acacatacac	acatgtgcat	accttagatc	acaaattcaa	89580
catctcaact	tagccgccaa	agccacnttc	tgtgtcttgt	ctattctgaa	tatgctttta	89640
ggttaaatta	gccatgaatc	tgttcagatc	agtcattcct	gaaaatgatg	atgatattag	89700
ccatccttta	ccgggcacct	tactcagtcg	ccggcattat	actaggctctg	caggtagacatt	89760
tgctcattta	acacctagct	cagtcctaatg	aggtggatgc	cattgtcatc	tccatttttat	89820
catttaataca	actcagcctt	cctctctcct	gttccccact	ctgcttctag	ctgccactca	89880
cctgactctc	ctcttgatct	gggctgacgt	ggctgagtgg	gtttagcaga	actcttgtgt	89940
cctctttact	tggtttgttt	caagttcatg	tcctgcttca	tcttttccct	aactcactgc	90000
aagggtgaatt	ttcttctctg	caatcgga	ttccccaaat	cacatcaggt	cattaattta	90060
ttcactcatt	tgtattcatt	cattcattag	ttcatccatc	catctattca	ttcattcatc	90120
caatttactg	agtgcctcct	ctgtgccagg	cactgctggc	tttggaatta	aaaaccctgg	90180
ccctcaagga	atttctattc	ttctgggagc	aggatattta	catgggtgac	tgcnaagcttg	90240
atctgtgtgg	tactagtggg	aaacatagga	tggtgggcac	aaaagagtca	caaaacgccg	90300
ggcgcggtgg	ctcatgcctc	taatcctagc	actttgggag	gccaaggcga	gtggatcacc	90360
tgaggtcagg	agttcaagac	cagcctggcc	aacatgatga	aaccccaact	ctactaaaaa	90420
tacaaaaagg	aataatagct	gggcgtgatg	gcgggtgcct	gtaatcccaa	caactcggga	90480
ggctgaggca	ggagaattgc	ttgaacctgg	ggctttggag	gttgcaatga	gccaagagcg	90540
ggccacttca	ctccagcctg	ggtgaaaggg	tgaacctctg	tattaaaaaa	aaaaaaagtc	90600
acaaaaaagg	gctggccacc	taacctcagct	tgaggatggg	aaggccaggg	caggcttccct	90660
ggaggagggtg	actcatgggc	catgtcatga	ggatggggta	agaggagaag	gtatgtttca	90720
gacatctggc	tttcttaata	ttcttggtct	tcacccattt	ttcctgccag	caacataggg	90780
gagaagactg	agaccagcag	atacaaagcc	actgtactct	gtcctcacc	cctccttttt	90840
tctcccttc	ctactaccca	gatctgaggt	tttgagaaat	ctcttctcta	attatctgct	90900

tggtcattca	tttgttgaga	atctactatg	tgcatagcat	tatgaccaca	aaatgtcctg	90960
gtcgtttgtca	tcatgagagg	tcttaccocct	tcctctccag	ccacttctag	ggatttttga	91020
ctctgtccctt	ttccagaact	tggtctccagt	ctgggttgctc	gccatgaagc	acttacagat	91080
aaacctcatc	ttgggccagt	gcttccattt	actgtctcct	tttggcttgc	ttatccttcc	91140
ttctgccttc	ttgaattgat	ttgcctcttt	atctccctt	tcagccttga	aagttcctcg	91200
aaggcaggga	ctgtgtcccc	atcttttcta	taaatggcat	tggtgtacat	tggttaaagtt	91260
gagtgaaaaa	atctttttga	taatcacaaa	ttattaccat	atattgagca	cctaccatga	91320
tgagtatctc	tgatcttcat	atctttgaaa	tggtgatact	attaccttca	cttaattggg	91380
aagaaaatth	gtgctcaggg	aggtaaagtc	actttctcac	agtcacacag	ctatttcaca	91440
ggagagtgtca	gtatcaaatt	taggtttttc	tggtcccaaa	ccctgatgtt	tccccacat	91500
cattgtttct	cagacttggc	tgcaatcacc	tgaaaaatth	tattaaattc	tttttttttt	91560
ttttttttga	gacagagtct	cgctctgtca	cccaggctgg	agtgtcaatgg	tgccatctct	91620
gctcagtgtca	accactgcct	cccacgttca	agtgtatctc	ctgttccagc	ctcccaagta	91680
gctgggatta	caggcaccca	ccaccacact	cagctaattt	tgatatgttt	agtagagact	91740
gggtttcacc	aggctgatct	caaactcctg	acctcaagtg	atctgcccac	ctcaccagc	91800
caggttttat	taaattctga	tgcttgggtc	ttaccctagc	agttctcatt	agttgtctca	91860
tggtgtggcc	tggtgtttgg	gatttttgtg	aagccccag	ctgagtctaa	cgggcagtca	91920
tgtttgagaa	ccagtatact	acaccatgtt	gccttctctg	tttcacacag	ggttggtgcta	91980
gtgtgtgagt	ctaagcctac	atgggtggat	atccacgtga	tcaggctgtca	agttctttgt	92040
agtggaggat	ctttggccct	cctgtctgct	tcccaacca	gcgaacctac	cacaccatgc	92100
acgtgcacag	ccaagggttg	ttgactgttt	aatcagtctc	tggttcttta	tgtctcatcc	92160
gggggattca	ctgataggag	gctggagctt	gatcacatcc	agatgttctc	atcttcagag	92220
cctgttaaca	taacagaagt	cttataaaca	tgctgagcac	actcactggg	ctggagagtg	92280
ttgcaggatg	acctaggccc	attgcagggg	tggtctgggt	cgccctccca	cacttctgcc	92340
accctgcttg	ggaaggaggg	agctgttttg	aagttccctg	gtgcaatatg	atgtattgct	92400
ctggctctgc	tcaggaggaa	gctatggccc	acctagtcat	tgagggttag	ctaatacgtg	92460
tattctgttt	ctgttgaggt	ttcacagagg	ccttttttga	cccttcattt	tatagataag	92520
gcagtggagg	cacaactaca	tgaaatgact	cgacaaataa	atggatttag	aaccaggtt	92580
tctgactccc	agggtgggtg	tttttccatg	gtgtgacagt	gattaatgtc	taccttttca	92640
caccagtcct	caactgaaga	caccagctta	caccttcctt	cctgtttctc	ccaagaacag	92700
aaagtgaccc	cggatgtcgc	tttctgttcc	tggaaggcca	gttccagtgg	ttagaagtcc	92760
tgttcactcc	tggtgtgtgg	cctggggatg	gtcctgacat	ccctgggctc	ttcctggacc	92820
tggtccagcta	aaaggaaatc	tcctatgatg	gtactcagat	acttttgga	ccttgctcagc	92880
cctaattccca	tctcctagt	tttagtattc	agctaccctt	cactgggcag	taattctgtg	92940
ccaggcctga	tctgggcatt	ggtggtacta	agaacgcata	ttccctatcc	tataggcata	93000
gtctggtaga	gacaacatgc	aagtaaaaca	atgttcctta	aactgtgggt	ccacacctcc	93060
ctcccccaac	attaaaagt	taagggtatg	ttattcaaat	gtagatttgt	aggctctgtca	93120
ctctagaccc	actatttcag	aatctctggg	gactgggccc	aagaaactgc	attttcgc	93180
gctccctaaa	tgaagcttag	gtgctctgag	gtttgacaac	tgtagtagag	agcctaagtc	93240
taacagtgtg	gagtcacatg	tgatgggaag	catcaggtag	gtagcagttt	gcaggagcac	93300
tgattctgag	ggacactaac	tggtgcctaa	aacagctact	ggctgtcatg	aggaataact	93360
aggagctagc	catagagggg	tagcagtga	tcatttctct	agcgatgtaa	atcttgctca	93420
atthattctg	tctatataac	tcaatattac	tgaagtttgc	ctaaagcaga	atacacctgg	93480
atcatacagc	atthattgaga	gactggctgg	gctgtcaggg	cctcctgtta	ctttatctct	93540
gcatgtgacc	ctcttasstc	cgcggtattaa	ctcctgtcct	cattaagcct	cacactgtag	93600
ccccattttc	agatcaaaac	tgtttctctt	ctggtaaatg	atthcagttt	gcaaagtttg	93660
ccctctagag	gttgcttagt	gcctggccat	gtgggtctag	ttcatgtggg	cctgatgagc	93720
tggtttttatc	tttattacaa	agaagtttag	ctgttaggag	agtgggttgg	aaggagaaga	93780
ggtagacagc	caaatgagat	gagtcaggga	aaactatact	gtttcggagt	catagggctc	93840
ctaccaagca	tctggtcaga	aacctctcat	tttgagatc	aagaaattga	ggttcagaaa	93900
gatgacatga	ggtacgcagg	gacgccacca	gacacagcct	ccaactctag	aaactaaaat	93960
tctggattct	tagtgctgct	ttttctgttt	tggtgactg	gattgaagcc	ttttctaact	94020
gtactcagag	ggcctattat	ttagggagat	tccgtatgaa	atccttttagc	aatcaaatca	94080
tttaaatagg	ctgatgggtt	aaatatatgt	taatgtgttt	tcctaaagcc	tggtcagacct	94140

gggtttttga	gcttttgtatc	acatgttttta	tgttttggaa	gaaaatgaga	ccatgtctgt	94200
gaaggcactt	tgatatgcgt	aatgcactct	gccagtgttt	gtcaaaacat	ggtccccagg	94260
tcagcagcat	cagcatcacc	tgtaagtgat	tctccagtcg	catccccggc	caggmsmagg	94320
ccmrctcny	aggcytgtg	smmggcsmgg	nnsngnnntg	nccggcccg	cccggcccg	94380
cccggcccg	cagtctgcat	tttaacaagc	tctccaggtg	attctgatgc	atacttaagt	94440
ttgagaacca	ttgcttgttt	tgcattaaac	aggagattag	tctctgcagc	ttgtgggaat	94500
aaagctttaa	atctctccaa	tttttagctct	gtgaaaaggc	agtggggaga	caggaatgaa	94560
cggactagt	ccacaaaagct	caggtgggg	gggtgagatc	atttagaaga	gaaagaccg	94620
gcatggtggc	tcacgcctgt	actgtcagca	ctttgggagg	ccaaggcagg	ttggatcaca	94680
aggtcaggag	tttgagacca	gcctgcctat	catggtgaaa	ccctgtctgt	actaaagata	94740
aaaaaaaaa	aatttgccag	tcattggtgat	gcatacctgt	aatcccagct	actcgggagg	94800
ctgaggcagg	agaatctctt	gaaccggga	ggcgggggtt	gcagtgagct	gagattccac	94860
cattgcactc	caacctaggt	gacagggtga	gactccgtct	caaaataaaa	aaaaaaaaaag	94920
aaaaggaaa	gctgtgtgtg	tgtgtatgtg	tgtgtgtgtg	tgtgtgtgtg	tgtgtaacag	94980
caccatcaca	ctgtttgagt	tgaggagcac	atgctgagtg	tggtcaaca	tgttaccaga	95040
aagcaatatt	ttcatgcctc	tctgatatg	gcatgctcc	cctatctcat	tctgtgtgt	95100
gtttagccag	gcaactgttg	atcatcaata	ttatgataac	gtttctccac	tgtcccattg	95160
tgcccacttt	tttttttttt	ttgagttact	tactaaataa	aaataaaaaca	ctatttctca	95220
atagacttga	agcttcaaga	tttctgggtg	gacaatgaaa	ccttctctgg	gttctctgat	95280
cacaacctct	ctctcccaaa	gtctactgtg	gacaagatgc	tgagggtga	tgtcattctc	95340
cacaaggtaa	gctgatgcct	ccagcttctc	cagtagggtc	gatggcaatt	acgttgtgca	95400
gctactggaa	agaaatgaat	aaacccttgt	ccttgtaatg	gtggtgaagg	ggaggagggt	95460
agtttgaata	caacttcact	taattttact	tccctattca	ggcaggaatt	gccaaaccat	95520
ccaggagtgg	aatatgcaac	ctggcgctcat	gggcagctg	gttaaaataa	aattgatttc	95580
tggtctatca	cttggcattt	gtgatgattt	cctcctacaa	gggatacatt	ttaggttgag	95640
ttaaacttaa	aaaatattca	cagttctgag	gcaataaccg	tggttaagg	ttattgatct	95700
ggaggagctc	tgtctaaaaa	attgaggaca	ggagacttta	gacaagggtg	tatttgagga	95760
cttttaagaa	ttttataaaa	taagggtctg	acgcagtggc	actgagttga	gaactgttgc	95820
ttgctttgca	ttaaatagga	gatcagtcct	tgcagcttgt	gggaataagg	ctttaaatct	95880
ctccaatttt	agctctgtga	gatggcactg	gggaaacaga	aatgaacgga	ctagtgtcac	95940
aaagctcagg	tgggatggac	gagatcactt	caaaggtctg	taatcccacg	tctataatcc	96000
cagcactttg	ggaggccaag	gcgggaaaat	cacttgaggt	caggagtctg	agaccatcct	96060
ggccaacatg	gcaaagcctg	tctctactaa	aaatatgaaa	attagctcag	cgtggtggca	96120
tgctcctgta	gtcccagcta	ctcgtgaggc	tgagacagga	gaatcgtttg	aacctgggag	96180
gcgagggttg	cagtgagcca	atatcacgcc	attgcactcc	agcctggctg	acagagttag	96240
actccatctc	aaaaaaaaaa	aaaaaaaaaag	aattttataa	aatcaggaaa	taatattagt	96300
gtttatgttg	aatttttaact	ttagaatcat	agaaaacttc	ctctggcatc	attattagac	96360
agctcttgtg	cagtgggtag	caccagaccc	agcttgcatg	gttattgatt	tttcagagac	96420
actttttgag	cttattctct	ggcagaaagg	ggaactgctt	cctcccctat	ctcgtgtctg	96480
catactagct	tgtctttaca	agaagcagaa	gtagtggaaa	tgtttattct	tgaaaataag	96540
ctttttgctt	cacatgatct	agaattttta	aaattagaaa	aatgtgctta	ctgctgccc	96600
ttctgaaact	aggaaaatat	gccttgtgtt	taccaattgt	gtggttagga	gatgggcca	96660
aggcatcagg	cttttgaaag	tagttgcatt	tagcataatt	tccattgccc	cctgccaatt	96720
tcatatctgt	cacatcta	cagtttaaaa	taaggggcat	cctaagcatg	gagatgggtc	96780
ttgatgtgtc	cttgagttt	ctgtattttc	agtattcttt	tttttgagca	tacaagacat	96840
ttattgaaaa	attcttggga	tcaatacttg	tgtaaggaaa	ggaaaggga	caaagcatga	96900
ttgggcagag	gcagaagatg	acatcaacaa	aggccctccg	ttgtcagtat	tctttttttt	96960
tttttttttt	tgagttggag	tctcgtctctg	tcaccagggc	tggagtgcag	tggcgtgatc	97020
tcagctcact	acaacctctg	cctcccaggt	tcaagtgatt	ctctgcttca	gcctcctgag	97080
tagctgggat	tacaggcatg	caccaccaca	cctggcta	tggtgtattt	ttagtagaga	97140
cggggtttca	ccatgttggt	tggccagact	ggtcttgaac	tcctgacctc	aggtgatccg	97200
tctgccttgg	cctcccaaag	tgttggaatt	acagacatga	cccactgcac	ccggcctgtt	97260
gtcagtattc	ttaaacatag	acactaactg	taggctgaca	gcctagcagc	aaggaccagt	97320
taaagaaatg	agtagaactg	aagtgtgctt	gagtatctct	ggcagtcagc	aaaaacttaa	97380

tgggaatcat	ggtaggccaa	atgttctgca	gtatttcaaa	agctgcatgg	gttttgagag	97440
gcttttggtc	catcactcac	cttaggttgt	tctacaagca	catcagcctg	ccccaattta	97500
aacagggcag	ttagtaaatg	tgtaataaag	agtctgcatt	gttcattcct	tcaacaaata	97560
ctggctataa	gttttcagca	ctgtggatgc	aaagtgcagc	ggataaacag	gctcttcttt	97620
caaagcttgt	ggtccactgg	accacgtatg	aagtagaata	gttttaggtcc	agaaaggcaa	97680
ttaagtaaaa	tatgaccaag	aagaggctct	ctagtgggtt	tggtataaag	aaaagataag	97740
aatgatttag	aattggccta	tcaatgagat	aagaggcctg	gcttttctggc	actctgctct	97800
agggcaagta	aaatggagaa	ttccaaattc	tgaaattggt	agaacatagt	tctgtgtctt	97860
agttaaatat	ctacacttac	agataaatag	cataaatgct	ttctcccat	atttcagccc	97920
agtctactt	aaagacaaca	taaattgcaa	aatagtgcag	atgttggttca	tctaataaaa	97980
gtggttccag	gaattcagac	tctggattcc	tgtttgccaa	atcatgtgtc	ccactcttaa	98040
gaaaacgagt	tggactctgg	atttttcttt	gcaagaggga	caagagtgtg	ggagatactg	98100
agttaatgca	acttgcaggt	tttaagtgtc	ctgtcattgt	gccttgtgct	ttgatacatt	98160
ctgagtttca	gtaaagagac	ctgatgcatt	ggactgttgc	aatggaacct	gttttaagat	98220
cttcaaagct	gtattgatat	gaagtctctc	aaaagacttc	aaggaccagg	cttccaatct	98280
tcataatcct	cttgtgcttg	tctctctttg	catgaaatgc	ttccagggtat	ttttgcaagg	98340
ctaccagtta	catttgacaa	gtctgtgcaa	tggatcaaaa	tcagaagaga	tgattcaact	98400
tgggtgaccaa	gaagtttctg	agccttgtgg	cctaccaagg	gagaaactgg	ctgcagcaga	98460
gcgagtactt	cgttccaaca	tggacatcct	gaagccaatc	ctggtgagta	gacttgctca	98520
ctggagaaac	ttcaagcact	aatgctttcg	gaatgtgagg	cttttccttg	gacagcatga	98580
ctttgttttg	tagaaaagta	cggctggctg	ggagtgtgtg	atataattta	gttcagtggg	98640
attctaagtg	ttcttagtgt	tctttcagac	ttttgggcca	tctcccaaag	ggtgaatggg	98700
aagaataagc	tgggtgtggc	tgagttaaag	ccaaaagtgt	tttgtgcttg	tttcaatcag	98760
agaagacctg	ctttttcatg	tttttactat	tataatacta	agcaagagct	catttgaaaa	98820
cagagttctt	catatttaaa	aaaaaaaaagt	cttgaaacca	ttgatgggaa	gatggatatc	98880
tatttatgtt	taaaaaacca	tcataaagat	gacattgtgg	gctgtcacag	ttggaaggcc	98940
ctggaattag	atgagaccac	actattttagc	ttacttagta	ataacattgc	aaagaaaaat	99000
tccgacgaag	ttttttcagc	ctaggaatca	atagttcaga	gaagcactct	atgagaatac	99060
ccattcattc	ttaacaaaaa	aatactggtg	agcctgagca	gtttgggtcat	cagagtgttt	99120
tatatagtgc	cagaacaaat	atgtctctag	gtgttctgag	agctctgggtg	aaattcctct	99180
cgctacccca	aacatcatca	tttaatatcc	aggattctgg	ttttctactc	accagataga	99240
ttctcttaaa	accagggaaa	gattcctgga	ggaaggatgt	atctggaaaag	agatgttcct	99300
tattataata	aaatgaaatt	gtaatactct	tggattttgt	gcagcacgaa	ttcttttatag	99360
agagttgggtc	ctcccagaga	attaagaata	ctcagtttct	ggaccctggt	cccagatcat	99420
accctagaat	gtgaccttag	aaacacactt	caggattcat	acctttgatt	gaccatcaaa	99480
aagtttttgt	atcggccagg	tgtggtggct	cacgcctgta	atccccacc	tttcggatgc	99540
cacggcgggc	agatcacgtg	aggtcagcag	tttgagacca	gcctggccaa	catggtgaaa	99600
ccctgtctct	acaaaaatac	aaaaaatagc	cgggcgtgat	ggtgggcacc	tgtaattcca	99660
getactcggg	aggctgaggc	aggaggatcg	cttgagccta	ggaggaggag	gctgcagtga	99720
getgagatct	gtctcactct	gttgcccagg	ctggagtga	gtggcgcgat	ctncgactca	99780
ctgcaacctc	cgtctctcat	gcttaagtga	ttctcatgcc	tcagcctccg	agtacctggg	99840
actacaggca	cctgbcacca	cgcccagcta	ttttttgtat	ttntagtag	acatagggtt	99900
tcactatggt	gcccagctgg	tctcgaactc	ctgagctcaa	gtgatctgcc	cacctcgggc	99960
ctcccaaagt	gttgggatta	aaggcatgag	tcaccgtgcc	tgggtcccatg	ttataatttt	100020
aaagtaaggt	atatttctct	acagggatct	ttgcaacctt	aagtaanctg	gcctaaaaag	100080
ttagagaagc	tgacttgtgc	agacatttgc	agcctgttgg	tcttttttgt	gctgtgaatc	100140
atagagggtg	aaaggttatt	atgaatggta	caaaactttg	ttacaaaacc	attttcttgg	100200
actgttttgg	gotgcttcac	tgcatgacaa	atgctcacc	tttcagctgg	aatgattgaa	100260
attttggaaa	agatgggtgt	ttttagaaga	cattgttaatt	tggtccgggtg	ctgtgcccct	100320
tcattccatt	tcacttctgt	ttactcatta	aacacctatt	gtgtacacaa	cccggtaaaa	100380
tccctccact	cacacaatgc	ctgaattata	ctcatagtag	aatgactgtt	tagccctcat	100440
catctgataa	ttaacagctc	aggtttcaac	ctgacagtat	ctctctggga	ggattagcag	100500
cgtgacagag	tgcagggaaa	tgcaccttca	gaaccgtcag	ctacactgtg	tcccatcctg	100560
ctgtgttgtg	gttgtgcctt	gtggatgcgt	tggtttatga	ccaggtattg	attaagggtg	100620

ctactaccag	gtgctttctg	catatctcgg	gtttgtggag	cactcagggt	ctgcttctgc	100680
ccctctgctg	ttaccaagag	acctctcttc	aaaatggggc	tcttgagtta	gagtagaatg	100740
agtgatcagg	attgttttgt	gtaaagatga	tttctgagga	aggctttagg	atgaaatgac	100800
ttccaaacat	tttgaaatgt	gactcttact	tattgaatta	agcagggcct	taattggaat	100860
gctgggactg	atacttgatt	tgcattaagc	agcctttttc	tattgctgct	tggnttgaaa	100920
tttcaacatt	tgtgatggta	gatggatgtg	acatgtgatg	acattgcaca	tgggcagtta	100980
actgtgccaa	gaagtgcagc	agtagcagca	accggagatg	caaagcccaa	catgatgggg	101040
agagaaaactc	ttctttcaat	atgtgcttct	gtaccaaag	tggaatttca	cgagagacat	101100
attttgaaac	atttctcctt	ttgtgtgtgc	gtgagtgttt	ccctgtttcc	agccaagggt	101160
attgtgagtt	tctcctgggc	ctccttcaga	atctgggtgc	tctggaaagc	agtgttttgg	101220
caacatgggg	aaagtatggc	agtgtgggag	ggtcagctgg	gtctgggttt	gaatattgca	101280
tttgaatatt	ttaccagcat	tgatgtcggg	taaattat	agtccctgta	agcctcagtt	101340
ttctcttctt	ctacatacac	ataatatatt	tgactctttg	ttgtgattat	tggttacaca	101400
tatgaagagc	ctggtgtggg	gcctggcaca	caatagggtg	tcaataaata	gaagttgata	101460
atttaattga	catgagtagt	agaaattatg	tccttgaaaa	caattgcgtc	aagatagaag	101520
ttttcagcca	ggcacagtgg	ctcacatctg	ttgtaatccc	agcatattgt	gggggcccag	101580
gcggatgaat	cacttgaggc	caggagtcca	agaccagcct	ggccaacgtg	gtgaaatccc	101640
ctctctacta	aaaatacac	tatttgccag	gcaggcgtgg	tggcgcacac	ctgtaatccc	101700
agctactgaa	gaggctgagg	cacaagaatc	gcttgaaccc	aggaggtgga	ggttgcagtg	101760
agctgagatc	actccactgc	attccagcca	gcgtgacaga	gtgagactct	gtctcagaaa	101820
aagaaaaaaa	agatagaagt	tttcttctgt	agatcagtg	tagaactcat	accaagcgaa	101880
gtggctcctg	tgagtatttc	agtgaataac	tgcattcttg	ctcagatatt	gtcaagactt	101940
ttcacccaaa	gattcttatt	tatgtctcag	tccgtacctt	gtgtgaaaat	taatactgga	102000
tgtcagaacg	ctggtgtggt	tttaaagtcc	cctgggggta	agagcagttt	ccattaggtg	102060
ttctctgctt	tttacttaaa	aatcttactc	atgcattgag	caatatttat	tcagttctta	102120
ttatgtgtca	ggtattttct	aggagctgga	ctcaactcaa	aagatatact	tttgatgaga	102180
acaaaggtgg	gtggatatat	gaaatattat	ctgtgggata	aatgcactta	gtcatgaggg	102240
agacttggtta	tggagtgcgc	tcattgtatt	tgtactgttg	agttaacaac	ttctaggagg	102300
agctcagggc	cacctggcag	gggcttcttt	tgtcttgctg	ctcagcaagg	tgtattttgc	102360
tgtagagtgt	gctgggcagg	tgaacttttc	ttacttttct	cttgggtcct	tcctaaagca	102420
gcatgtacct	ttcccagagc	gaggagaggg	ccaccttctt	gtctcacaga	aacctccaat	102480
ctgttttgga	ctgcaggaag	gagccatagt	agtggaccag	caaatttttg	cccagagagat	102540
tgatttgctt	ccgattgtac	tttttttttt	tattgctact	ataacaagtc	accaaatact	102600
tagcagttta	taaccacata	catttattat	ttcataaggc	cagaagccag	gatacagtag	102660
agctcaccta	ggtctacttc	ttagaggctc	acagggtgta	tatcaagggtg	tttgtggggc	102720
tgtgtttttt	tgggggggct	ctgaagatga	atctgctttg	gagctcatcc	aggatatcag	102780
atgaattcag	tcccgtgcag	ttgtaggact	gaagtccatt	tccttgctgg	ctgtcagcag	102840
gggctgggtt	ttgctcctaa	agggtgccat	cattccttct	tatgcttttc	atgtgactcc	102900
tctcccacaa	gtgggttgag	tatctctcac	actttaaaac	tctgacgcct	cctgtcacat	102960
ctgtotccag	tcagagaaaag	ttctttactt	ttaagggtcc	aggtgtttag	tttggggccc	103020
ttcacataat	ccaggataat	ctccttattt	tgagggtcat	aactttaatt	acatctacac	103080
agtgcatttg	ctatgaaatg	tgcattttca	cagactccag	gaatagggtg	tggacatctt	103140
tgcaggggat	attcagtcctg	ctgtagcttt	cttgattgga	aggaatagtt	tatcatatat	103200
ttgaagtgtt	ccctcagtcg	ctttggcttt	tgttgacccc	tctgcagctc	tctgcttttt	103260
cttgcccata	tttggaaggt	gactctaaag	cataataagc	atcacctatt	agggttttta	103320
atgtacaaac	caaacagagg	tgactttggg	aggagaacat	ctctgaacta	ggtatgagac	103380
attcatcgaa	aaaaatccat	caagtgttta	ttgtatgtct	gctttatacc	agcactgttc	103440
taggcaactga	agttcgacca	ccaacaaggc	agttgtgatg	cacttggagc	tttcattctg	103500
atgggcattg	agatccaggc	tgaaggctga	gtctgggaat	ttgaggaatt	ctttgtaggt	103560
cctgggtcta	cagagtgaga	gctgtcctcg	ttccagtttc	actgatgacc	tctcgaccag	103620
ctccctcaca	gcagtctttg	ccaacacaga	cactgtgggc	tgtagtggga	ggaatgtggg	103680
gttgaatgag	ctaggttttg	gctctgtcct	tgactcacca	ttgcctcagt	gatgtgaaaa	103740
tgggtgtctga	tccttaagggt	tgggaagtcca	ggcatgcaga	tttatctcca	tctcaataac	103800
gtggggaaaa	aaaagaagtg	gtttttgtgg	aaccagtggt	tccactgtcc	acgtgtcttc	103860

cagtgtgcct	agcacattcc	caccaagtgg	atcttgggtc	atgagggaaa	gaagggaaa	103920
tgaggggtgcc	ccgaggctcc	cagaagacaa	gtaatgtcac	agctgagctg	tgtacagatc	103980
gaagaagcag	atggataagg	agtgaacaaa	gtcatccttg	tcttgagggg	tctttatttta	104040
gcttcttctt	gactcttaga	caaggaccca	gaatacagat	ggggcttggt	gttaccttca	104100
gcctcatggc	tttttagggc	tagatacttc	agcttggtac	atgcagtcct	taaagcgtct	104160
gtgtgggttt	ttgcaggaga	gaacacttgc	tcgtgttctc	ctgtctggaa	cccggacatt	104220
gttggaagt	atcagatttt	gtttggcttt	gtgtgatttg	actgcccacc	tgtttacttg	104280
ctttctccca	gcaagcagcc	gcaatcccca	tgggttggtg	atgggtggaa	tgacacactg	104340
tgtagattta	ctcttcagac	tctatgttca	cctcattctt	atgggaaaag	aagagcacta	104400
gctggtagat	atagcagtgt	gttaatatga	ctgctcaact	ccattttacac	agacattttc	104460
accttagtta	tactatttct	tcattaaata	ttgttgccag	atctaagata	caggtttaaat	104520
tttttctct	gaattatgtg	gtagtagatg	tatttaacta	tgtttagaaa	cagcaaaaaa	104580
tgaagcgttt	gaatgcgtta	aacacatcta	atttgaaagt	taatatattag	gttggttgatt	104640
tattttttta	aagattagaa	tattccttag	aaatgtagtc	ttataatttc	gtattttcaat	104700
aaaaaatatt	aaaatgtttc	ccagaggaaa	tctttactgt	catagaaatc	accagaaaga	104760
gatagcaatt	actcctgggt	ggtgatagtc	tttgatttgt	ggtttacttg	ttttcagttt	104820
gaactaaaac	atactaattc	cagcctgagt	tggattttgc	atatggtcaa	agtgaagtag	104880
agatttttgt	ctactttatg	aagatattta	aaggacattt	gaaatgtttc	aatgaacaca	104940
ttgtaacatg	cattcccagg	aggaaaacaa	aattgtgtat	tgtgttgaaa	atactatttc	105000
aatatatgta	taccacgctc	tcattttgtc	atagaattcc	tgaaaaattt	agatgtagat	105060
gagattttta	taagttgaaa	atattttcag	ttgcatttta	attgcacaga	tgtgttttct	105120
acttctattt	ggctgtgact	cctaatgcat	tgtaatatta	tcttccagtg	ttcatttctg	105180
gtagtataaa	tggattaaca	aatgcattca	ttcatttagt	aaacattttac	tgagccctaa	105240
tcagtgccat	gctctggccc	agggaggtag	tcagtgtcaa	ggtccagccc	tctctgtttc	105300
atttctgggt	ggggaatggg	tgggacagac	caggaaaagg	gcagctgccg	tatagagcag	105360
gagacgctct	gctgggtgtg	atacagtcct	tgatggggcc	tcctactgca	gtctttccag	105420
tcaggaaaaga	aacatcctct	aagtggggac	ctaaagaatg	gatgtcagcc	aggtaaagaa	105480
cagggaagtag	aatgtttccag	gtagaggaca	tagtatatgc	attggtccag	agaggagagg	105540
gaagagaagg	tggcatttag	gggaactgta	cctgattgag	atagttcagc	atggagcaga	105600
tatagctgtg	gggagtttga	tgggagaagg	aacgatgaga	cgtgatgcct	gcagagggtg	105660
gcaaaggctg	gctgagggcag	gtctttcttc	acgagccatg	tgaggatata	gggccttatc	105720
ctcagagtaa	agggagccat	taaaggctcc	acaggctgag	tgacttgatg	aggtgaatac	105780
gaaacagagc	atttctgtga	atgtgaatgg	cctttggagc	agaactaagt	tcatgaggat	105840
ggaaagaata	taatgaagcc	atctcttaga	cccagaaaaga	atggggcaca	acagccttta	105900
tctttctggg	gccaatgtca	cagaatgcca	tgcttttagg	aaacatggta	tgttgtgatt	105960
aacacatttt	gcagaagtgg	gtgggagcct	ttgaataata	acagtaagca	tttgtgcatt	106020
cttctgttaa	tgacattaca	gttatgatct	gaaaatattg	agtcatacat	gaattcctgt	106080
tatcttaact	cagaaaatat	agtcctcac	taaaggtttt	attttccttc	tttttcccat	106140
ttcctttact	tcgtataaga	aagtcacttg	tctcctgggt	gcaatggaga	cctatgtgag	106200
ttcatagcca	agagaatggt	tttggttaga	aaaataatag	taggaattcc	aagctgtgaa	106260
ttttttactg	aagctctttg	gaaataggat	ttggcaagtt	ttgtctgcct	tcgtcaagta	106320
agcatgagca	ggagagcaca	gttaatagca	ggtgcagaca	catgattctc	agaccgtatt	106380
ttgtgttcta	gtttcaaggc	atgaattctt	tcctgggggt	aattttattc	aagggaagtta	106440
tctgtctgtt	agatctgata	tgtgctcagg	ccaacataga	ttctttaccc	ttcctttctt	106500
cctgctcacc	tgtccttctt	cttttatctt	ttcattgaat	taaaaagaaa	attatgaaat	106560
agttttcaaca	tgaaaaaagg	tacagagaat	aacataaaga	acactcctgg	ctgggtgtgg	106620
tggctcacgc	ctgcaatccc	agcactttgg	gagtctgagg	cagccagatc	actggaggtc	106680
aggagttaga	gaccagcttg	gccaacatgg	tgaacactg	cctctactga	aaatacaaga	106740
attagccagg	catggtggcg	tgcacttgta	atcccagcta	cgtgagagac	tgaggcagga	106800
gaattgcttg	aaccaggag	ggggagggtg	cagtgagctg	agatcacacc	actgcactct	106860
agcctgggtg	acagagttag	actccgtctt	caaacaaaca	aacaaacaaa	aagaacactc	106920
ctgtaccatc	atccatcatt	ttgccgtgct	gactccaggt	tctattttaag	aaataaaaaca	106980
ttacaggtag	agctgatgcc	acctctgttt	ccctagctca	ttcttcagag	ataactcttg	107040
tcttgtagtt	ggatgtttta	atcctctata	tcantgtata	cttacattct	atgtataaca	107100

atatttggta	ctggcctaaa	tgtgttcaca	ttgtataagt	gtgcatattg	gcctgccact	107160
tcatttggaa	ttatgttctt	gagatttatc	aatgttgata	catgtggaat	ctgggttaatt	107220
tttgccatag	tattctatct	ttatactaaa	cttttaaaaa	tccatgcttc	tagtcttttg	107280
cttatttttt	caggttatgg	tatgttttgg	atgcacagaa	aagtaaaatt	aagtcattgag	107340
caaaatatct	ggataatcca	agcttttaaac	ttgatgtaga	atgtgaatca	tgtgtgtttt	107400
gttaaccctg	tgatgtcaat	ccatgcctga	ttgtgttaact	ccaaccaata	ttccttttgaa	107460
aatggaaatt	tgtttatatt	gactacagat	tgccaatatt	attagtaaat	gctgagcact	107520
taatctcgaa	taaagaacta	gtttaaaaat	gattctaaca	atggcattga	ctgttctacc	107580
ttattactca	tgggtgggtt	cagccaatgt	ttctgttggg	gacccaaaacc	aaaaacagtc	107640
aaattaaaca	agcagtcaaa	cccaacatac	agactactga	taagaaggtc	atatcataag	107700
atatggcatt	gaatttgtgt	ctgctaattg	aaaaatctga	tgcccacagc	aaacttaata	107760
aggacctatg	ttacatcca	tgctcaatta	cattcctggg	ttaaacagtc	atgcttttagg	107820
ccctgctgtg	tgcttgagg	tttgtggaag	tgtggggctt	ttaagagaag	gagaataagc	107880
ttgctccaga	gttaagaaat	ttaaactaaa	agtcctaaag	atgttggaaa	aactattgcc	107940
cttgaagatg	taaattcatt	aagtgtgaga	agacctttta	tacaaaacaac	agacccattc	108000
actgatttgt	acccttcagg	agacagatga	ccggtaattg	tgacaatggg	tgaatgttgg	108060
gtttgggggt	tttagaaaaca	tctgcacttg	gtgactactg	tatctaattg	gtgtgacaaa	108120
cctggcaccc	catgtgtttg	gcaccatctt	gggtcctact	cagggccagg	tgaaccgagt	108180
ggcctcttca	ctgcttcaga	gtctgaagca	gattgtagta	tgccggaccag	acacagaata	108240
taccaccaag	caggttggcg	caaagcatac	tgggaaggga	ggctttgtga	acatggtgct	108300
ggttctcaaa	cctcagtgtc	caaaagagtc	tcctgaggat	tcctggatca	cactcttaaa	108360
cttctcattc	agtaggtgtt	agctgggtga	gaatctgcat	tttttttttt	ttttctgaga	108420
ccgagtctca	ctctgatgcc	caggctggag	tgtattggcg	ccatcttggc	tcactgcaac	108480
ctctgcctcc	caggttcaag	caattttcct	gcctcagcct	ctctagtagc	tgggattaca	108540
agcacatgcc	accatgcctg	gctaattttt	gtacttttag	tagagagggg	gtttcgccgt	108600
gttggccagg	ctcgtcttga	actcctgacc	tcagggtgatc	cacccacttt	ggcttcccaa	108660
agtgttgaga	ttacagggtg	gagccaccat	gccagcctg	aatctgcata	tttaacaagc	108720
accacagggt	attctgatac	agtagctccc	caaacctcac	agtgttagtg	aatcccagtc	108780
atttacaatt	ctgccatgat	tttggtcata	ttcaagtgca	gctggtagca	tttttagtta	108840
atataatttt	taaattaagt	cacttctttt	ggataattaa	atttaattac	aaggggaagc	108900
ataccactgc	tgtaaaaaaca	tcacctgctt	taaagagaag	gtacataatg	aatatacatt	108960
aagataaaga	tgtatatgtg	tgtgtgtgtg	cacatataag	tatacacata	cctaccatag	109020
ggattgagtt	ttccttcagg	tttttcaact	gaaatgtcaa	ctttgaggcc	agttaatatg	109080
tgtaagatat	atgtgtgtat	gtatgtctat	acatatagac	atatacgtaa	aaacatacat	109140
ggatgcatat	agtatatcta	tacacaacct	attatgcata	tcatgtatat	ttcatccact	109200
tagtattatc	ttnttatttt	gccgtttggc	aaatgctcag	taaaagaaaa	gggttagaag	109260
gggagaaagg	catttttatcc	caagccttca	ggaatcagga	tgaggatgtc	ttcaccttgt	109320
ggtggggagt	anattataca	attagagaca	gcacattgga	gntgtggctg	atatgctgtg	109380
tgatgatagc	tctagctctc	tgccatagcag	aggaaggaca	tttcaataga	agaaaaagtt	109440
taagaccttg	ccgagaaaaca	gagaaaggat	gtttgtcttt	ttaagaagtt	gaaaaccctg	109500
tttgacagaca	aaagccctcc	agttttggca	gtaaactttc	atgcaaggga	agaaaaaggc	109560
aggggatgac	attgttgaca	attgtgagga	attaccatgt	gccaggcact	gtgcgagggg	109620
ctttgtacat	atcctctagt	tttagtgctt	ataaaaaactc	tgtgatatgt	gcacagcatt	109680
ttaaactttg	ctgcatagtc	gagaaaaatgg	aaggatgggg	aatttgagtc	atttgcccag	109740
ggttctatag	ctaccccagg	ttcccatgac	tggagaattg	gggcacaggg	tggcggggga	109800
gagtgagtga	caagaatcct	aacaatctta	tttccattga	gtccttataa	agaagtgga	109860
ttactacca	cgtttttaag	tttttcttaa	atttaggtta	tgtggatctg	gcgtttcttg	109920
ttttgtcctg	ggtttgtttt	gtttttgcta	tgctgtcttg	aacatctgtc	atctttagtg	109980
cctaacggta	aacacaaaaa	cactttacct	cctatagctt	tcaattaaga	tctctcagtt	110040
tgtgtttgta	atagttttcc	aggcaagttc	tccttaggtt	cggcttctag	tgtgttaacc	110100
tttagttata	aagtgaacc	aaagagagaa	agtagaaaca	aaacacctca	cctgtttttg	110160
ctcatgaatt	actctctatg	gaaggaacaa	tcatgaacac	ctctgcgtat	cacagaggcc	110220
tatctgagtc	tgacgtttaa	gggagaccgc	gtaggtccct	ttgaggactg	tgaatgtggg	110280
agtcctggga	ctctggtgaa	gaacccgttc	cagaagagat	gaatgagctg	gacaagttct	110340

ttcatagaac	cttttaggcag	gttttcttag	aaatgcacat	tgaggattat	gcttggatat	110400
tgtgatgatc	agaatgatac	tcaatccctt	ctgcatttgg	aattctcttt	gaaagaaaac	110460
atcccaggca	gctattttctc	agagatagtg	agtcccagcc	acttcttagac	attttcttgt	110520
gtagtctaca	ttataatttc	acagcagtc	ctgatatgac	aaatgtcaaa	atagcccaac	110580
cttctctaaa	cttcagagat	gtctgatatg	atattgaata	aaacaatgct	catagaaaca	110640
tcaagaaagg	tggattttcc	ctggatactt	ttttcctgct	tgacaaataa	cagtgaagaa	110700
actgatctca	cgtctttttc	tctttggaag	cctgaacact	cagaacccaa	cttgaggctc	110760
ctcagctata	gcaattctga	cttcacagtc	tgtaaattat	tgttcttttt	tttcttttagc	110820
ttatgctttc	tgccttaatt	tatcttttcc	ctgttcta	gaattattgt	cctatatctg	110880
ctgtgcagtt	aggtgacata	taacagcaat	taaatatatg	aattggtaca	tataaagatt	110940
tgactaaaac	tcgatgtaaa	aataagtgtt	ctacattcaa	tttccagtg	tagaaacagt	111000
gctgacttga	acagagtgc	agaattccat	ctttccctat	ttttgacagc	tttaaacttt	111060
atattttctt	cctttcttgt	gagccgtcat	taactttgtt	tctcaaagnc	cattcccgt	111120
ttacccatct	tgcagacgca	gacagatttg	ggaatttgcg	gtcagagttg	tattggacac	111180
atccccccag	cccacatgag	atccttttaa	tctattgcat	attaactagt	tttaagtaca	111240
atattcctac	ttcatttaaa	accattaatc	aaagaatgag	tttgaaaatg	aacaaaatgc	111300
aaacttacag	ttagaaataa	ttgtagtgtc	tttagttttg	gttaggagtc	ggtttcttgt	111360
ttgttaaact	caagattgtg	aacagtttta	attcacttgt	ttatttccaa	tagagatttc	111420
aggtttacat	ttgaattcag	aaacaaagtt	ttctttctca	ttacagagaa	cactaaactc	111480
tacatctccc	ttcccagca	aggagctggc	tgaagccaca	aaaacattgc	tgcatagtct	111540
tgggactctg	gcccaggagg	taagtgtgtg	ctttccagta	ccaggaagcg	gatcatccac	111600
tgtatcagta	ttttcattcc	tgagtctggc	aagaggtcct	tttgagttga	atatcacatg	111660
ggatgtaata	tcaattttca	aagtataagt	gatgtaaaca	ataatgtttt	gatttcctta	111720
ttttagaaat	gaagaaacct	aaaactcata	gatgtctcag	agctaattgg	ttagtggcta	111780
acagctggat	atctagttta	gaaccttctc	cattttttct	ttttgcccct	aggtaatcat	111840
acatttgtaa	agaggagaat	tatctctgcc	actgcccatg	cactgctttt	gtctgaccag	111900
caatttctcc	atattgcttc	ttcagtagca	aggccaatca	ttttaccaac	acacatgctt	111960
gctaactaac	aggaataacg	tggtagccct	aattcagccc	tttcccttga	aagcatctgg	112020
cttctgaggt	tcaactatgg	gaatatggtc	tcttaatgaa	cattaagttg	agtttgctt	112080
ttaggtccac	atgttgacaa	atgtatcaga	gtaatctctg	tcctaggatc	agagggcctg	112140
taggcacttg	caaaagcagt	tagctctgac	tcccagccag	tgcacactcc	acctttctga	112200
ctcccagcct	tgtctcaaat	taggcttgga	agcgaggaa	tgtctggtgt	ccccagcat	112260
aggaagctga	gccagggggc	agtgtctaca	aacaatacag	actttaacgt	gtaggatatt	112320
ggaaaataat	aatttgtggg	gaaattgtct	cagacttgg	ccacccttat	ttttagctgc	112380
ttctctaata	cgtttttctt	tttttggtgc	ttgtatctaa	cctaccatt	ttttggtgct	112440
tgcatacttt	tttcaaatat	caaaaacgaa	ctttatgttt	tctaacaatg	aaagtattgc	112500
atgttcattg	tggaaaatgc	tgaagacttg	gaaaatacaa	aatgctgag	atcaaacact	112560
attgatacgt	tagtgatttt	cttctgtctc	tgttctactt	tctttctttg	aattctgctc	112620
acgtgtttct	gactgatgag	gtctgacttt	tgggttctct	ttccagagga	gaagccttct	112680
ttcagcttgc	catttgttac	cctggttatg	aaggctggta	acctttttta	ctaggtagag	112740
aagctggacc	aactgggggt	cttccagggg	gagaatgaga	aagagaaact	gttttgcaag	112800
tcogtagcta	tttctctagg	gccctgttag	ctgacattga	catgccttgc	attgctctgc	112860
agatccccct	gcagccctct	gtcccttggt	catttctggc	cttagagaaa	gcaaagcagg	112920
gtctgtaaca	ggggaggctg	cctctaaact	cagggtttgg	ttacagctgt	tttcacttac	112980
atcactggcc	ctggtttttt	ttttttttct	ggcattaaaa	aaaaaaattg	gaagcagggtg	113040
atgtttccat	tgtgatgtg	gtggaaactc	tccaagtga	caatatacgt	ttttcttggc	113100
agctgtttct	tgtgcccctg	ttgtccttgg	tccaggacaa	gcaaggacca	tctgcctctt	113160
tcaatagaac	acctccagat	ccctttgatc	aaaagttact	cattgtctga	cttgctattt	113220
ctgtgagata	aatgggagaa	gatcaataaa	tgcacttggt	tgtccagtca	gcngtgtgga	113280
aagttgataa	ttttgaccaa	agcacaacct	ttgaaaggaa	aagaaaaagg	gagtgaatgt	113340
cttctgagaa	gctgcctagg	ttcagacagt	gtcaccatt	tccctgtatg	ctccacatga	113400
caaacctgag	tgggtctcat	catgtccatt	ttgcagatgg	caccaaggct	cagaaaaggtt	113460
aggcaacttt	tcagtcacc	caatgagtta	attgacaaaa	ctgggattca	aaccagaaac	113520
tgttggattc	caaagcctgt	gttgttgctt	gcttcgtgaa	aaactccagt	agcgactgga	113580

atagaaagga	gaaccttcca	agaaagaaaa	tacgcactag	cagaacctgg	aaattgggag	113640
gaaatgagga	cttgaggaat	aagatgaatg	aaagctgacc	tgagtttcac	atctgggtga	113700
tgggaagggg	ggacagggag	gcagcatctc	agatgtccac	ccagcaccga	ccagctgcct	113760
ggcattgcta	ggtgttgagg	actcagcagt	gaacacgcta	acttctctgc	tttcttgggg	113820
cacgtatagg	gtgagagaca	gaaacaaaca	ggtcagtgtg	caatgccaca	ggagggatat	113880
atgcagttaa	gaaaaagcag	ggtaaggggc	atagagcatg	agaaggtgct	ttttttaaag	113940
gggktgatta	ggaaaagctct	ctctaagggtg	acagttggac	ctgaaggaga	tgatagcatg	114000
tctgtggtga	gggaaggaaa	ctccgaacag	gaagaatggc	agatacaaag	acattgatgc	114060
tagagcatgc	ctaagggaatg	tgtttaagga	ccagggaaaag	tgagcaagtg	gtgggggggag	114120
gagaggagct	cagagcagga	ggagggtgagt	gccatacagg	cctggcaaga	ctttggattc	114180
ctgctgggtg	agatgagaat	ccagcggagg	gcttgaggga	ggggacatga	tgtgatctag	114240
agtttagact	gtttacactc	tggttggttg	gttgagaaga	gactgggatg	ggggaaaagg	114300
aggacaaagg	acattgtgct	ggattgagaa	agcagtaagt	cagtttcatt	cattcactca	114360
accgatgatg	ttcaaatacc	accatcatcc	gtgggctaaa	ggatgaagag	ccatccctcc	114420
ctgagagtca	ggaagcactt	cccagataaa	gtttggagtg	tgagctgagg	tgtaggagaa	114480
agagtaagag	tttacccttg	aaacgggtgc	tgggaagagt	caatagtttg	gaataactca	114540
ataatttatg	gtgcttcttt	agaaaagattt	gctggcttta	tgtgggaaga	aatttktttt	114600
tttgattggg	gagtgggtggg	ttgggtggtga	ggctgcctgt	ggaaagagaa	gtgagtgttt	114660
tgactcactg	ttattttaaaa	atctctagggt	ctgttccaat	aagcaacaaa	aggcaaaatg	114720
gcctggttct	ctgtccccctt	tctgtctgta	tgctcgtac	aggttatgaa	aagaaaaagt	114780
tgggaaaagc	tgtccacctc	acctaattgt	gttcttgtgg	agtgtgctag	atgccccctc	114840
tctggagaaa	aaaaatcctt	gtggcctctg	accacactct	ggagagccta	gttccccctt	114900
ggaggcagaa	ggcaaagctt	aggacctaga	gagtgtcgtg	ccacgccact	cacaggaacc	114960
agcaggctgt	gaggttgaaa	gctaggcata	tggagctttc	caggctgggt	gcagggcctc	115020
gtggcccttc	ccctccccctc	tgtgtcttat	agctcagtct	tcccaggcgg	tgtgaacacg	115080
cagtgcattt	tccaggaata	cagggattta	ttaatgattt	cttgtgaaat	gtttggaaat	115140
acaaagtact	ctataaatat	ttcataatag	cattggggct	gagaactcca	caaagtgccg	115200
gaatacatatt	gcatgtaaga	cagaacgctg	cctgggtcat	tgatgcctgt	tgagtggcag	115260
tcacagacac	tgcttagggt	ttctgactca	cgctgttggg	actgttctat	gcagggcacc	115320
ctcttgtgtg	gcataggatt	tgtgcctcac	cacacactgt	tgtagctttg	ctgtcttgat	115380
gatgagtaga	gggcagtgtc	caggccatgg	tataagcatc	tactgcccc	cagggttacc	115440
aaaaccaagc	caagtttgtgt	ctcagcgagc	tccgtgaagc	atggagaagt	tgagtactca	115500
gagacatgac	gtgacttttc	aaaggctgta	agctgacgag	ggacatagct	agggttcaga	115560
cttgagtttt	tctttttctt	tttctttttc	tttttttttt	aagactgagt	cttgctttttg	115620
tgcgccaggc	tggattgcag	tgggtgcttg	ctcactgcaa	cctctgcctc	ccgggttcaa	115680
gcaattctcc	tgcttcagcc	tccccagtag	ctgggattac	aggcacctgc	caccatgcct	115740
ggccaacatt	tttgtatttt	tttagtagag	atggggtttc	accatgttgg	ccaggctggg	115800
cttgaactcc	tgacctcagg	tgatccaccc	gcctcgacct	cccaaagtac	tgggattaca	115860
ggtgtgagcc	actgcacccg	gcccagactc	gagtttttca	tcttaatgct	ttttcattgc	115920
ctgacacttt	actgagacca	agatagggaa	cttcacatac	agtacctttt	ctcccaaggc	115980
ggaagagggc	tgttcaattt	ctacactaga	gttcggggag	ttttagaaat	gagtcagtta	116040
tgcaggatga	gagcagttcc	tgataggctc	aaccacaatg	agatgtagct	gttcagagaa	116100
agcattcttt	tatctataaa	ctggaagata	atcccgggtga	aacgaagccc	agccccagg	116160
gcttcaactaa	ctccaggctg	tgcttctcaa	actttagtga	gcataggaat	cacctgggca	116220
tcttgtgaag	ctgtagattt	gaattctgca	ggtcggcaga	ggggtctcag	aatccgcatt	116280
tccaacaatg	tctccagtaa	tgctgatgct	gctcgtccct	ggaccacaga	ttgggtagcc	116340
aggttcttgg	aagctcatcc	caaggctttg	agatgacatc	agacaaaata	tgttctggga	116400
catggctttt	gagaggtcaa	gaaaataaga	tgtttctttc	tcttctcatc	cccaaccctt	116460
gcaactgcct	tttctccctt	ccccaccct	cctttctgtc	cccatccctg	acgccagctg	116520
ttcagcatga	gaagctggag	tgacatgcga	caggaggtga	tgtttctgac	caatgtgaac	116580
agctccagct	cctccaccca	aatctaccag	gctgtgtctc	gtattgtctg	cgggcatccc	116640
gagggagggg	ggctgaagat	caagtctctc	aactggatatg	aggacaacaa	ctacaaagcc	116700
ctctttggag	gcaatggcac	tgaggaagat	gctgaaacct	tctatgacaa	ctctacaagt	116760
gagtgtccat	gcagacccca	gccctgtccc	caaccccatc	cctcccttag	ttctggcctt	116820

ggcctgtgtc	atctcctccc	tctgtagcag	cgtagatgt	ctacatgccc	atttgcccac	116880
cagactgagc	tcttcctaga	ggagagaggc	ttctcttgaa	tagctacctg	tccccagttc	116940
tctgaatgca	gcttggcaca	tctcaggtgc	acagtagtgt	ttatcaatgg	aatgaatgat	117000
tgacagccaa	ccttctggtt	ttctggggga	tgtggaaggg	tggcttccag	ggtgatcaag	117060
aatgagataa	tggcagaagg	acaaatcctg	caagatctca	cttatatatg	gaatatatgt	117120
aaggtagaaa	gtgtcagttt	cacatgatga	ataagttcct	gggatcttga	tgtacatcgt	117180
gatgactata	gttagtaaca	ctgtatatga	tacttgaaat	ttgctaagag	agtagatccg	117240
aagtgttcac	actacacaaa	aaaggcaact	atgaggtgat	ggatttatta	acagcttgat	117300
tgtggtgatc	cttttacaaa	gtatacatat	attaaaacat	cacattgtat	accttaaata	117360
tatacaattt	ttatttgtca	gttgtaactc	aaaaaagcta	gaaaagcatt	tttaaaaagg	117420
atgatgtact	ggtcttaata	ttaccattga	gataagcttt	ataataacat	aaaaagaaat	117480
aacagtaatg	ataatagcaa	caacaacaac	aacaaagaac	taacatttaa	gtagaatttc	117540
ttgtgcaactg	tgcattctgt	ttaaagttatc	tcattttacc	ctcatgataa	ccctgcaggg	117600
aagattcttt	aaccccat	ttcataggct	cagagaggtt	aagtgccttg	gttagagcca	117660
catcagagtt	aatccacaag	agccaggatt	caagcccaaa	tctgcctgga	tctgtgctct	117720
ctaagataac	tgtagtggt	ggcgtgtgtg	ttctcacact	cagacatttg	atctgccctt	117780
tgtttcccat	tcttagctgc	aaggcagtgt	ttaaagaaccc	tgtgtctcca	tatccactcc	117840
ccacacttaa	gcacttttgt	gggcccgtgt	gccgtatgcc	tcgtggcagc	agggatccaa	117900
tgtcacagtt	ttaggcagtg	gcaccccttt	ccttgaaaac	ttgatgcagg	ggaacctttc	117960
tccatttcca	accacaggtg	tgtcnntttc	agacactgag	tnagggcagg	ttttgtactt	118020
tattgtaaca	caagaacctt	ttcttctctg	gagtaaagca	ctccagacat	tcgcaagttg	118080
ctttacaagc	cttaaaagga	tggtattgta	ggcaactttt	attaaatccc	atctcctcct	118140
ctccccagc	ttgcaagttg	acccaaggaa	gccttcattt	ccatgacaga	cttaattgtg	118200
agggcatcct	cattaaaaaa	aaaaaaattc	tattatcttt	ccagcatata	gaagatactt	118260
ggtatctaaa	aatccctgaa	aaacttagaa	tgaattttta	aaaatcaggg	atcctgctgg	118320
ataaccaaac	ccatttgtct	gttacaactt	ttgtatttgg	gtttttgtta	agtgtacata	118380
tactagtttg	tgtaatttaa	agagaatttt	tttttttttt	tttgagaggg	agtctcgctc	118440
cgttgcccag	gctggagtgc	agtggcgcca	tctcggtcca	ctgcaagctc	cgctcccggg	118500
gttcacgcca	ttctcctgcc	tcagcctcct	gagtagctgg	aactgcaggt	gcccgcacc	118560
atgccagct	aatttttttt	ttttttgtat	tttgagtaga	gacggggttt	cactgtgtta	118620
gccaggatgg	tctcgatctc	ctgacctcgt	gatccgcctg	tctcggcctc	ccaaagtgct	118680
gggattacag	gcgtgaacca	ctgcgcctg	ttgagaattt	tttttttttt	ttttgggaga	118740
cagagtttcg	ctcttgttgc	ccgggctaga	gtgcagtgc	acaatctcgg	ctcactgcaa	118800
cctctgcctc	ctgggttcaa	gcgattctcc	tgccctcagc	tcatgcgtca	ccacgcccag	118860
ctaattttgt	attttttagta	gagacaggg	ttctccatgt	tggtcaggct	ggtctcgaac	118920
tcccaacctc	aggtggttcg	cccgccttgg	cctcccaaag	tgctgggatt	gcaggcatga	118980
gccactgcgc	ccagcccaaa	attttggttt	ttgcttgaaa	actgaggtct	gaattcagcc	119040
ttctggttgc	ccctcaagag	tcagtttaaa	tgttggtcat	gttagttgtc	agtgaataca	119100
atggtgaggg	tggcatgaga	gtgtgaatct	ggatgggagg	gcttgtgctt	catgaataca	119160
tttttccaga	tcagctcagt	cgtgagttat	ccgtcattga	cgttataata	agctctgatt	119220
atztatcaag	catcattctt	tatagatatc	tcagtttaat	ctgagataat	cttctccaca	119280
tctctccaca	tagatgttat	gaattttact	tttacagagg	agccaactga	ggctcagata	119340
agttacttat	tatatgacta	gtagtggtag	agctgggggt	tcaactaaga	actctctggc	119400
tccaaagccc	ttgtaagttt	ctatcagtat	atgacctgc	atatgagcat	ttgtctctcc	119460
tcttcttcat	agctccttac	tgcaatgatt	tgatgaagaa	tttgagagt	agtcctcttt	119520
cccgcattat	ctggaaagct	ctgaagccgc	tgctcgttgg	gaagatcctg	tatacacctg	119580
acactccagc	cacaaggcag	gtcatggctg	aggtaagctg	ccccagccc	aagactccct	119640
ccccagaatc	tccccagaac	tgggggcaaa	aaactcaagg	tagcttcaga	ggtgtgcgct	119700
aagtatactc	acggctcttc	tgggaattccc	agagtgaata	cctcaagtct	gatgcagacc	119760
agagctgggc	cagctcccca	gtcgtgggta	tagaatcata	gttacaagca	ggcatttctt	119820
ggggatgggg	aggactggca	cagggctgct	gtgatggggt	atcttttcag	ggaggagcca	119880
aacgctcatt	gtctgtgctt	ctcctccttt	ttctgcggtc	cctggctccc	cacctgactc	119940
caggtgaaca	agaccttcca	ggaactggct	gtgttccatg	atctggaagg	catgtgggag	120000
gaactcagcc	ccaagatctg	gaccttcatg	gagaacagcc	aagaaatgga	ccttgtccgg	120060

gtgagtgtcc	ctccccattat	taccatgtgc	ctgcttgata	ctggagaggt	gagtttctgg	120120
tcacttttccc	aggtgtgagt	gaggtgagaa	ttcttttcagt	ttatctagct	gggggaatgt	120180
agttagcata	gctaaagtca	cagggcacca	cctctccaga	agtacaggcc	atgggtgcaga	120240
gataacgctg	tgcataatcag	catccatgcc	actcacggtc	aaatagcagt	tttctgcaaa	120300
acttagtgag	ggctggtgtt	tggaaagtga	gttgagtaat	tgcagtaccc	tattttcctt	120360
tttngctgca	gcctctcagc	cagccacagc	atctccctgt	gtcttggtag	gttttgga	120420
gaagtgtggg	agcaaaagca	tgatgttaca	tgtagactgg	cctgagatac	tcattctcag	120480
ggcactgtgt	gaatgatgag	ctgctgttac	tgtgtggagg	ggaaatgcac	ttagtgtctt	120540
agagccactt	gaaagggata	agtgtcttag	agacaattgg	gttcaaagt	ggagcaggct	120600
gagcaagaac	agaatgtctc	ctttgcctga	gcctgagtgc	tgtaaatcac	atcttcctgc	120660
cttgggctga	gttagagaat	cattagacta	tttctgttt	ccatgggtgag	ggaggcctct	120720
tccttttgtc	tctgtctccc	ttaagaagca	ggtgaggatt	ttgccagggt	tcttggtttg	120780
aaccttattg	actttaaggg	cggctgggtt	ttagagactg	tacctaccta	gggggaacac	120840
ttccgaagtt	taggactatt	ccctgatccg	ctgggaggca	ggttactgag	gaagtccctt	120900
taaaaacaaa	ggagtttata	ctgagaaaag	cataaacagt	gatttgtatg	gattcacact	120960
gactaatata	gctcatgcc	ttaaagtggg	gtctcttctc	taaaggaggg	ttatatgatc	121020
tagccccgta	gacctaagt	tggtttcaga	cctgttcttc	ctggctctct	ccttggaatc	121080
catatttcta	ctagttaggac	tttttctgtt	tgtctggctc	tcagaggatt	ataggaggcc	121140
ctgtgaagt	actcagtga	ttttgatttg	tgggcaagta	gatggttccc	tagtctgaaa	121200
ttgactttgc	cttaggtgct	tcaattcttc	ataagctccc	agttcttaaa	ggacaagatc	121260
cttgtaaaca	tggcaatggc	attcattagg	aatctagctg	ggaaaatcca	gtgtgtatgc	121320
ttggaaatga	gggatctggg	gctggagaga	aaggcatggg	catgccttgg	agggacttgt	121380
gtgtcaagct	gaggaccttt	actttaagct	ctaggggacc	aggcaagggg	agatgtagat	121440
acgttactct	gatggggtgg	atgaattgaa	gaaggatgag	gcaagaatga	aggcagagac	121500
cagggaggag	gctctccaag	tggccaaggc	ataaagcaag	aaatgaggcc	tgggtgactgc	121560
ttagtggcag	agcagtgaaa	gagagggagg	catcaaaagt	agtctcgatt	tctagctggg	121620
tgggtggtag	cgatgtccag	tagggcagtg	gctactgagg	tctgcagtg	aggagggtgg	121680
ttgggctgga	gacagatgat	gagggagtca	tcagcctgtg	ggtggaagaa	aagggaacct	121740
cttccaactg	ttttctttgc	ttcttccctc	tctttctctt	tttttttttt	tttgacaga	121800
gtcttgcctc	gtcaccacag	ctgaaatgca	gtggcatgat	cttggctcac	cacagcctcc	121860
gcctcctggg	ttcaagcaat	tctcctgtct	cagcctccag	agtagctggg	attacaggca	121920
catatcactg	tgcccggtca	atttttgtat	tttcagtgg	gatgggattt	caccatgttg	121980
gtcgggctgg	aatgaactcc	tgacctcaag	tgatccacct	gcctcagcct	cccaaagtgt	122040
tgggattaca	ggcatgagcc	accgcgccc	gccttcttct	cctctcttaa	agagtgttta	122100
tttaattcca	caaacatgag	cttgtcaccc	cctgtagcct	ggcatctcct	acacgagggtg	122160
atggctgagg	cttctgcttc	tgctggggta	gctctgatct	ttctgctttc	tctggcactg	122220
tctacccatg	ttgcctcacc	ccacaggctc	cagggcacct	ctctcgggca	agtcttgga	122280
ccctctgaca	ctgatttgct	ctcttttctg	agctgctttt	agccacccat	cctcgggacc	122340
tgttttctct	ctgcctccac	ccctgcgggc	agtcttaggt	ctcctgcccc	tcacgagcac	122400
cccagagagg	ccacgtgctc	agtgatctca	gtgggcgc	ctttctagtc	ttgctattct	122460
ttttggccat	gttgttcaga	aaccatactg	ggcagggccg	acttcaccct	aaaggctgcg	122520
tctcttcaact	ctgcttttgt	ttgttccaaa	taaagtggct	tcagaattgc	taaccctagc	122580
ctctgtgaac	ttgtgaggta	caattttgtg	tctgttatgt	taacaaaaat	acatacatat	122640
cttctggtg	atggtataaa	ttgctattct	ctattggaaa	gcaatttgga	atgaaaaattt	122700
aaagaacct	tttaaatat	gctatcctgc	gtacctccat	tccacccacc	cccagggatg	122760
tagcctactg	aaataatttt	aaagaagtca	ccatatgaga	gaaaatgtta	ttgctatatt	122820
gttattgtga	gaaattggaa	atagactaaa	tgttcagcac	tataggaata	attaatgaaa	122880
ttacatatac	tctatacaat	cattatgctg	ccattgaaat	aataaataca	aaggcgcaag	122940
gggggaaaaag	cttataatgt	tagtgaaact	aagactgatt	tttttataaa	gcagcagttt	123000
tcagaccctt	ggagactcca	attcggtaga	accagagctt	catcttctct	gtcgaagctg	123060
tgacaggagt	tgcaaatgcc	tctccttttt	gctgagtttg	cagctgctgt	ttttccggca	123120
gcacatctgt	gcaggcctct	gcctcggccc	ctctggnatc	tgctgattga	gcagcggatt	123180
gatctgtcct	tctctttcgt	gttgacccat	gtgaggaacc	aactggcaag	ggaanacaaga	123240
aatggaaata	ggcctccttt	gcatcatgac	ctgtacatcc	tgcaattgga	aaagattgta	123300

ctttagttgg	tttaaccagc	agcattatatt	ttctaaacta	agcagtaaga	aggaattagg	123360
ttttatgtgg	gatcaacaga	ctgggtctca	aaagaggaag	gtgatagaac	acagtgggga	123420
gggggaggtg	cactagaaac	agagggccta	tgctttcatt	ctggctttgc	tacttaatag	123480
ctgtgtgacc	caatcttaga	gacttaacct	ctctgaactt	ccattttctc	atgtataaaa	123540
tgggaaatat	taaaggatac	tcactgggct	ggtggttgt	gcctgtaatc	ccagcacttg	123600
gggaggttga	ggtgggagga	tcacttgagc	ccaggtgttc	aagaccagcc	caggcaacat	123660
ggcaagactc	tgtctctatg	aaaaaattaa	aaattagcca	ggtgtggtgg	tgtgcacctg	123720
tagtcttagc	tacttggtag	gctgagatgg	gaggatcact	tgggcttggg	aggtcaaggc	123780
tgcggtgagc	tgtgattcca	tcactgcact	ccagcccggg	cggcagagcg	agacactgaa	123840
tccaaacgac	aacaacaaca	aaaggcaaaa	aaataaaagt	gccctcttta	tggagttgtg	123900
taaggtgaag	catatacact	attcaacata	gtaactatat	aaaggaagta	ttgttgttgt	123960
tactgtagtt	aataaccatta	agtgaatgt	ttcgtatagt	ggaaagcaca	tggactctga	124020
attcagactg	gtctgacttt	gagctcagc	tccacatcta	gtaatactat	gaccaagccc	124080
tggttaaaat	catgtttttt	tttcttcagc	ctcagtcctc	tcacatataa	aatagggaca	124140
ctgtcattta	cctcagtttt	ctgtgaggat	aaaacaacga	cagtgtatat	gcaagtattt	124200
tgtaaatttt	gtagtgtctc	tcaagattta	gttggtgttt	actacttgta	ctttctcact	124260
ggaatggcag	atgctgttgg	acagcagggg	caatgaccac	ttttgggaac	agcagttgga	124320
tggcttagat	tggacagccc	aagacatcgt	ggcgtttttg	gccaaagcacc	cagaggatgt	124380
ccagtccagt	aatggttctg	tgtacacctg	gagagaagct	ttcaacgaga	ctaaccaggc	124440
aatccggacc	atatctcgct	tcattggagg	gaatctgttg	ctgggatcat	ttagaaaaga	124500
cttaacggct	tctttctctg	agacgttaca	ataaggttca	ggcaggaggc	aagtttagaa	124560
ataatgtata	gtctcattta	caaaactatc	cctcaagcct	aacacaggat	ttgataacaa	124620
aaggcactta	ataaatgtta	gttgagtgg	tgaatgagta	aataaactct	agcttttagta	124680
aattaaactct	agcttattct	atataggctc	aagagaatat	ttctacccat	tttcttctag	124740
gttttcttat	ctcagtgact	aatggtagca	aagcattccc	ttaaaaaggc	attattttgtg	124800
aaacttatct	aaaatcgaat	tcgggtccaa	ttaaattttt	gaaattttat	attaaaaatt	124860
aatattagtag	ggatgggttaa	gaggtgtttt	ggtctggttg	gttggttagt	tgctatgact	124920
cagaattgct	aagaaaacag	aaaagtaaga	taagatcatt	gttttaacct	cttttctctc	124980
acaaaatcaa	taaataacat	atccctaaat	tactcttaga	atttctctta	aattgcagtg	125040
aaaaaccaa	atccttcatt	cttggttgaa	ggttggaata	ctacgttaga	gaggattaga	125100
gagagaggat	gagcaatcgt	gtagtacgac	cttgctctct	agtgtaggat	ttgtctcagc	125160
cactgcttgt	tgtcctggct	gccaacgttc	tcattgaaggc	tgttcttcta	tcagtgtgtc	125220
aacctgaaca	agctagaacc	catagcaaca	gaagtctggc	tcattcaacaa	gtccatggag	125280
ctgctggatg	agaggaagtt	ctgggctgg	attgtgttca	ctggaattac	tccaggcagc	125340
attgagctgc	cccatcatgt	caagtacaag	atccgaatgg	acattgacaa	tgtggagagg	125400
acaaataaaa	tcaaggatgg	gtaagtggaa	tcccatcaca	ccagcctggt	cttggggagg	125460
tccagagcac	ctattatatt	aggacaagag	gtactttatt	ttaaactaaa	atttggtaga	125520
aatttcaaca	acaacaaaaa	aactcaactt	ggtgtcatga	ttttggtgaa	attggtacat	125580
gacttgctgg	aagggtttttc	ataggtcata	aaataacagt	atcttttgat	ttagcatttc	125640
tactcaaggg	aattaatttc	aggaattttg	gtggcaggca	cctgtaatcc	cagctactcg	125700
ggaggctgag	gcaggagaat	tgcttgaacc	caggaggcag	aggttgcagt	gagctaagat	125760
cgcacatttg	cactcccgc	tgggcaataa	gagtgaact	ccatctcaaa	aaaaaaaaaa	125820
gatacaaaaa	tagaaaaagg	ggcttggtaa	gggtagtagg	gttttgggca	attttttttt	125880
tttttttttt	ttttattgta	tggttctaaa	ggaatggttg	attacctgtg	gtttggtttt	125940
aggtactggg	accctggtcc	tcgagctgac	ccctttgagg	acatgaggta	cgtctggggg	126000
ggcttcgcct	acttgaggga	tgtggtggag	caggcaatca	tcagggtgct	gacgggcacc	126060
gagaagaaaa	ctggtgtcta	tatgcaacag	atgccctatc	cctgttacgt	tgatgacatg	126120
taagttacct	gcaagccact	gtttttaacc	agttttatact	gtgccagatg	ggggtgtata	126180
tatgtgtgtg	catgtgcatg	catgtgtgaa	tgatctggaa	ataagatgcc	agatgtaagt	126240
tgtcaacagt	tgcagccaca	tgacagacat	agatatatgt	gcacacacta	gtaaacctct	126300
ttccttctca	tccatgggtg	ccacttttat	ctttttatatt	ttattttttt	ttttgagatg	126360
gagtctcgct	ctgacgcccc	ggctggagtg	cagtggctcg	atctcggctc	actgcaacct	126420
ttgcctcccg	ggttcaagct	attctcctgc	ctcagcctcc	acagtagctg	ggactacagg	126480
ctcatgctgc	cacgcccggc	tgactttttg	tatttttagta	gagacgaggt	ttcaccatgt	126540

taccaggct	agacttcaac	tcctgagctc	aggcaatcca	ccctccttgg	cctcccaaag	126600
tgctgggatt	acagggtgga	gccactgcac	ccagcccacc	actttaattt	tttacactct	126660
acccttttgg	tcaaaaatttg	ctcaatctgc	aagcttaaaa	tgtgtcatga	caaacacatg	126720
caagcacata	ctcacacata	gatgcagaaa	cagcgtctaa	acttataaaa	gcacagttta	126780
tgtaaatgtg	tgcacttctt	ctccctaggt	ggtaaaccac	atttcaaaac	aacccaaata	126840
aaactgaaca	aagcttcttc	ctcttagact	ttttagaaaa	tctttcagtg	ctgagtcact	126900
aagctgccaa	gttctcattg	tgggaactat	gcctttggat	gtaatgattt	cttctaagac	126960
aatgggcgga	ggtgtagtta	ttgcagacat	ctgaaatatg	taatgtttct	tccagattct	127020
ggaaattctc	ttattctctg	tgggtggtgg	tgggtggtgg	atgtgtgtgt	gtgtgtgtgt	127080
gtgtgtgtgt	gtgtgtgtgt	agggatcagg	atgcgggagg	agctgggttc	tgcttgattt	127140
ggttctctgt	tttgcatgga	atagtgtgtt	tccttgtagt	gctatctata	gcttttcaag	127200
gtcaccagaa	attatcctgt	ttttcacctt	ctaaacaatt	agctggaatt	tttcaaagga	127260
agactttttac	aaagaccctt	aagctaaggt	ttactctaga	aaggatgtct	taagacaggg	127320
cacaggagtt	cagaggcatt	aagagctggg	gcctgttggt	atgtagttag	tatgtgccta	127380
catggtaaag	ctttgacgtg	aacctcaagt	tcagggtcca	aaatctgtgt	gcctttttac	127440
tttgcacatc	tgcattttct	attctagctt	ggaatctgaa	acattgacaa	gagctgcctg	127500
aaatgtatgt	ctgtgggtgtg	attagagtta	cgataagcaa	gtcaatagtgt	agatgacctt	127560
ggagatgttg	aacttttgtg	agagaatgag	ttgttttttt	gttttgggtt	ttagtacttt	127620
aacataatct	accttttagtt	taagtatcgc	tcacagttac	ctagttagctg	aagcaagccc	127680
ccaaagaaat	ttggtttggc	aacactttgt	tagcctcggt	tttctctcta	cattgcattg	127740
ctcgtgaagc	attggatcat	acgtacattt	cagagtctag	agggcctgtc	cttctgtggc	127800
ccagatgtgg	tgctccctct	agcatgcagg	ctcagaggcc	ttggcccac	accctggctc	127860
acgtgtgtct	ttcttttctc	ccttgctcct	ccttgggggc	tccagctttc	tgcggtgat	127920
gagccggtca	atgcccctct	tcatgacgct	ggcctggatt	tactcagtggt	ctgtgatcat	127980
caagggcatc	gtgtatgaga	aggaggcacg	gctgaaaagag	accatgcgga	tcattggcct	128040
ggacaacagc	atcctctggg	ttagctgggt	cttagtagc	ctcattcctc	ttcttgtag	128100
cgttggcctg	ctagtgggtca	tcctgaaggt	aaggcagcct	cactcgctct	tcctgccag	128160
gaaactccga	aatagctcaa	cacgggctaa	gggaggagaa	gaagaaaaaa	aatccaagcc	128220
tctggtagag	aaggggtcat	acctgtcatt	tcctgcaatt	tcattccattt	atagttgggg	128280
aaagtgaggc	ccagagaggg	gcagtgactt	gccaaggtc	aaccagccg	ggtagcagct	128340
aagtaggatg	agagtgcagg	gttcatgctt	tccagataac	cacatgctca	actgtgccat	128400
gctgtctcat	tggtagtggt	tcattggcagc	atctgaaagc	tatttatattt	cttagatata	128460
ttgggtggcg	attcttctta	agtttctaa	aacaataatc	agaaggatat	atattgttgc	128520
aggttagact	gtctggaagc	agaggctgaa	atagagtttg	atgtatgggt	atttatgagg	128580
gctcaatacc	tatggaagag	atatggaaga	tgcaggattg	ggcagaggga	ggagttgaac	128640
tgtgatatag	ggccaacccc	gtggggcact	ctagagaata	tgcagcttgt	tggagttggt	128700
cttcacogag	ctgaaacatc	cagccctttg	tgctccccc	aggcctccct	cctgacacca	128760
cctacctcag	ccctctcaat	caatcactgg	atgtgggctg	ccctgggaag	gtcgtgcccc	128820
agggcctaca	tggtctctctg	ctgctgtgac	aaaccagag	ttgctgatgc	ctgaggccgt	128880
ctactgacag	ctgggcaaca	aggcttccct	gaatggggac	tctgggcagt	gcagttttgt	128940
gtctgaacca	tacattaata	tatttatatc	cgaattttct	ttctctgcaa	gcatttcata	129000
taaagacaca	tcaggtaaaa	ataaatgttt	ttgaagcaaa	aggagtacaa	agagataaga	129060
actaactaat	ttaatactag	ttaccatctg	ttacaaatag	ttcctactga	ttgccaaagga	129120
ctgttttaaac	acatcacatg	ggcttcttct	tctatcctca	ctaacccttt	taacagacaa	129180
ggaaatgagg	ctcaggaagg	tcaaggactt	tattgaggtt	ccacagtagg	atacagttct	129240
tgctaaaagc	aacccctccc	tcattgctctg	ttatctaact	gcaaggggaa	ggtcagtggtc	129300
agaggtagtg	gtcccatggt	tgggtgcataa	gagctgctct	gagacaactg	catgctgggtg	129360
ggtcctgcag	acatgtaccc	atcagccgga	gataggctca	aaatatccac	aagagtttgg	129420
atgattgtgg	gaatgcagaa	tccatggtga	tcaagaggga	aagtcaagtt	gcctggccat	129480
tttcccttggc	tttttagacag	aaaagttacg	tgggatatta	tctcccacag	ctcttctgtg	129540
gtgccaccag	tcatagtctt	tatataagga	gaaaccagtt	gaaattacct	attgaagaaa	129600
caaagagcaa	actcgccac	tgaatgctg	agaaagccct	ggactctgtt	gtattcataa	129660
ctctgccatt	atttttctgc	gtagttttgg	gtaagtcact	tatcttcttt	aggatggtaa	129720
tgatcagttg	cctcatcaga	aagatgaaca	gcattacgcc	tctgcattgt	ctctaactg	129780

agtaggaata	aaccctgtct	tttttctgta	gatcatacaa	gtgagtgcct	gggattgttg	129840
aggcagcaca	tttgatgtgt	ctcttccttc	ccagttagga	aacctgctgc	cctacagtga	129900
ttccagcgtg	gtgtttgtct	tcctgtccgt	gtttgtctgt	gtgacaatcc	tgcagtgcct	129960
cctgattagc	acactcttct	ccagagccaa	cctggcagca	gcctgtgggg	gcatcatcta	130020
cttcacgctg	tacctgccct	acgtcctgtg	tgtggcatgg	caggactacg	tgggcttcac	130080
actcaagatc	ttcgctgtga	gtacctctgg	cctttcttca	gtggctgtag	gcatttgacc	130140
ttcctttgga	gtccctgaat	aaaagcagca	agttgagaac	agaagatgat	tgtcttttcc	130200
aatgggacat	gaaccttnag	ctctagattc	taagctcttt	aagggttaagg	gcaagcattg	130260
tgttttatta	aattgtttac	ctttagtctt	ctcagtgaat	cctgggttgaa	ttgaattgaa	130320
tgggaattttt	ccgagagcca	gactgcatct	tgaactgggc	tggggataaa	tggcattgag	130380
gaatggcttc	aggcaacaga	tgccatctct	gccctttatc	ttccagctct	gttggctatg	130440
ttaagctcat	gacaaagcca	aggccacaaa	tagaactgaa	aactcttgat	gtcagagatg	130500
acctctcttg	tcttccttgt	gtccagtatg	gtgttttgct	tgagtaatgt	tttctgaact	130560
aagcacaact	gaggagcagg	tgcctcatcc	cacaaattcc	tgacttggac	acttccttcc	130620
ctcgtacaga	gcagggggat	atcttgagga	gtgtgtgagc	ccctacaagt	gcaagttgtc	130680
agatgtcccc	aggtcactta	tcaggaaagc	taagagtgc	tcataggatg	ctcctgttgc	130740
ctcagtctgg	gcttcatagg	catcagcagc	cccaaacagg	cacctctgat	cctgagccat	130800
ccttggctga	gcagggagcc	tcagaagact	gtgggtatgc	gcatgtgtgt	gggggaacag	130860
gattgctgag	ccttggggca	tctttggaaa	cataaagttt	taaaagtttt	atgcttcact	130920
gtatatgcat	ttctgaaatg	tttgtatata	atgagtgggt	acaaatggaa	tcattttata	130980
tgttacttgg	tagcccacca	ctnccctaaa	gggactctat	aggtaaatac	tacttctgca	131040
ccttatgatt	gatccatttt	gcaaattcaa	atttctccag	gtataattta	cactagaaga	131100
gatagaaaaa	tgagactgac	caggaaatgg	ataggtgact	ttgcctgttt	ctcacagagc	131160
ctgctgtctc	ctgtggcttt	tgggtttggc	tgtgagtact	ttgccctttt	tgaggagcag	131220
ggcattggag	tgcaagtggga	caacctgttt	gagagtcctg	tggaggaaga	tggcttcaat	131280
ctcaccactt	cggctcctat	gatgctgttt	gacaccttcc	tctatggggg	gatgacctgg	131340
tacattgagg	ctgtctttcc	aggtagactg	ctttgggcat	ctgtttggaa	aatatgactt	131400
ctagctgatg	tcctttcttt	gtgctagaat	ctctgcagtg	catgggcttc	cctgggaagt	131460
ggtttgggct	atagatctat	agtaaacaga	tagtccaagg	acaggcagct	gatgctgaaa	131520
gtacaattgt	cactacttgt	acagcacttg	tttcttgaaa	actgtgtgcc	aggcagcatn	131580
gcaaaatgtt	ttatacacat	tgcttcattt	aattctcaca	aggcntactn	ctgaagtagt	131640
tactataata	accagcaatt	ttcaaagtgg	agaactgtga	ctcaaagacg	ttaagtaacc	131700
agctttggtc	acacaactgt	taaatgttgg	tacgtggagg	tgaatccact	tcgggttacac	131760
tgggtcaata	agcccaggcg	aatcctccca	atgctcacc	aattctgtat	ttctgtgtcc	131820
tcagaggggg	tacaactagg	agaggttctg	tttcctgagt	acaggttgtt	aataattaa	131880
tatactagct	ctaaggcctg	cctgtgattt	aattagcatt	caataaaaaa	tcatgttgaa	131940
tttttcttta	gtacttcttt	cttaatatata	tacatcttct	tgaccaagtc	caagaggaac	132000
ctgcgttgga	cagttttcat	atgagatcaa	attctgagag	agcaagattt	aacctttttt	132060
ggttcacctt	ctgatcctcc	cctaaggagg	tatacatgaa	atatttatta	ctcctgcctg	132120
aacttctttc	attgaatatg	caattttgca	gcatgcagat	tctggattta	aattctgagt	132180
cttaactttac	tggctgaggg	accttgagata	ggctccttat	ccctcagttt	cctcatctct	132240
aaaatgggga	tggcacctgc	cccgtgggtt	gttgggaagg	cttacagagg	tgcagaatgt	132300
acgttgtaca	tagcaggttt	cagcaaattg	tagctccctc	tttccccaca	tccattcaaa	132360
tctgttcctt	ctccaaagga	tgtgtcaagg	aggaaatgga	cctggctggg	aaacctcag	132420
aatactggga	tgatgctgag	cttggctcat	acctgtgcct	tgctttcagg	ccagtacgga	132480
attoccaggc	cctgggtattt	tccttgcacc	aagtcctact	ggtttggcga	ggaaagtgat	132540
gagaagagcc	acctgtgttc	caaccagaag	agaatatcag	aaagtaagtg	ctgttgacct	132600
cctgctcttt	ctttaaccta	gtgctgctgc	ctctgctaac	tgttgggggc	aagcgatgtc	132660
tcctgccttt	ctaaaagact	gtgaaaccac	tccaggggca	gagaaatcac	atgcagtgtc	132720
cctttccaaa	tcctcccatg	ccatttatgt	ccaatgctgt	tgacctattg	ggagttcacg	132780
gtctcgatcc	ctgagggaca	ttttctttgt	tgtcttggct	tctagaagag	tatcttttac	132840
ttgccccctc	ccaaacacac	atttcatggg	ctcctaacaa	gctagaagaa	agaggtaaag	132900
acaagcgtga	ttgtggaacc	atagcctcgc	tgccctgcctg	tgacatgggtg	acctgtgtat	132960
cagcctgtgt	gggctgagac	caagtggcta	ccacagagct	cagcctatgc	ttcataatgt	133020

aatcattacc	cagatcccta	atcctctctt	ggctcttaac	tgcagacaga	gatgtccaca	133080
gctcatcaaa	ggctctgcct	tctgggttct	ttgtgcttag	agtggcttcc	taaatattta	133140
ataggctccct	tttctgccag	tctcttctgt	gcccattccc	tgattgccct	tggtaaaagt	133200
atgatgcccc	ttagtgtagc	acgcttgccct	gctgttcccta	atcatcttct	cctacctcct	133260
ctttacacct	agctcctgtt	tcagtcacct	agaaatgctc	acagtcgctg	gaatatgtca	133320
tgttcttcca	cacctccatg	cctttgtagg	tactgtttgc	tctcacagga	gaactttctc	133380
tctaacttgc	ctatcttctc	aactcctcct	ttctctccaa	gatctagttc	cggatcccct	133440
cccctgagca	tcctccttg	gttctcaggt	agtcagtcac	tctctgccct	gaacttccat	133500
ggcacgtgaa	agaaaatctt	tttattttaa	aacaattaca	gactcacaag	aagtaatata	133560
aattacatga	gggggttccc	ttaaaccttt	catccagttt	ccccaatggg	agcagcatgt	133620
gtaactgtag	aatagtatca	aaaccatgaa	attgacatag	gtacaattca	caaaccttct	133680
tcagatttca	ctagctttat	gtgcgctcat	ttgtgtgtgt	gtgtgogtat	ttagttctat	133740
gcaattttat	catgtgtgaa	ttcatgtaat	tactagctca	gtcaagctgc	agaaatatct	133800
cattgtcaca	aagctccttc	atgctacccc	ttaantggcc	acnagccnac	ctcccttctt	133860
cctcagttcc	tgacacctgt	caaccactaa	tgcgttccct	gtttttacag	ttttattatt	133920
tctagaatgt	tacataaatg	gaaccataca	gtaggtatcc	ttttgatact	ggcttttttt	133980
tttttttcac	tcagcagtat	tcccttagat	ctatccaaagt	tgtgtgtgtc	aacagttcat	134040
tcctcttcac	tgctgagtag	tgttccctgg	gaggggtgta	tcacagttcc	atggcatttt	134100
tagatgtatt	ttttaaacag	ctttcagcat	cctctatttt	aattgttcat	caagtccctt	134160
ttcccaatag	actctgaatg	ctcctttatc	atcgatttcc	catcaccaac	atcagtaccc	134220
aaataggccc	taaataaaca	tttatagcct	cctgcctgcc	tgagaaacca	gggtggacat	134280
ggagagaagg	cacttctgaa	agttcaagcg	cagtgcscctg	tgctccttaca	ctccactcct	134340
cagtgtcttc	tgtgggttca	tttctgtctt	ctctcctgtc	acagtctgca	tggaggagga	134400
accacccac	ttgaagctgg	gcgtgtccat	tcagaaacctg	gtaaaagtct	accgagatgg	134460
gatgaagggtg	gctgtcgatg	gcctggcact	gaatttttat	gagggccaga	tcacctcctt	134520
cctggggccac	aatggagcgg	ggaagacgac	caccatgtaa	gaagaggggtg	tggttcccg	134580
agaatcagcc	acaggagggg	tctgcagtag	agttagaaat	ttatacctta	ggaaaccatg	134640
ctgatccctg	ggccaaggga	aggagcacat	gaggagtgtc	cgaatgtgaa	catgttatct	134700
aatcatgagt	gtctttccac	gtgctagttt	gctagatgtt	atttcttcag	cctaaaacaa	134760
gctggggcct	cagatgacct	ttcccatgta	gttcacagaa	ttctgcagtg	gtcttggaac	134820
ctgcagccac	gaaaagatag	attacatatg	ttggagggag	ttggtaatc	ccagggaactc	134880
tgtctctaag	cagatgtgag	aagcacctgt	gagacgcaat	caagctgggc	agctggcctg	134940
attgccttcc	ctgcgacctc	aaggacctta	cagtgggtag	tatcaggagg	ggtcaggggc	135000
tgtaaagcac	cagcgtttagc	ctcagtggtc	tccagcacga	ttcctcaacc	attctaacca	135060
ttccaaaggg	tatatctttg	gggggtgaca	ttcttttctt	gttttctttt	taatcttttt	135120
ttaaaacata	gaattaatat	attatgagct	tttcagaaga	tttttaaaag	gcagtcagaa	135180
atcctactac	ctaacacaaa	aattgttttt	atctttgaat	aatatgttct	tgtttgtcca	135240
ttttccatgc	atgcgatgtt	aggcatacaa	aatacatttt	ttaaagaata	ctttcattgc	135300
aaattggaaa	cttcgtttta	aaaatgctca	tactaaaatt	ggcattttcta	acccataggc	135360
ccacttgtag	ttattttaccg	aagcaaaaag	acagctttgc	tttgtgtggg	tctggtaggg	135420
ttcattagaa	aggaatgggg	gcggtgggag	ggttgggtgt	ctgttctctc	tgagactga	135480
atggagcatc	tagagttaag	ggtaggtcaa	ccctgacttc	tgtacttcta	aatttttgtc	135540
ctcaggtcaa	tcttgaccgg	gttgttcccc	ccgacctcgg	gcaccgccta	catcctggga	135600
aaagacattc	gctctgagat	gagcaccatc	cggcagaacc	tgggggtctg	tccccagcat	135660
aacgtgctgt	ttgacatgtg	agtaccagca	gcacgttaag	aataggcctt	ttctggatgt	135720
gtgtgtgtca	tgccatcatg	ggaggagtgg	gacttaagca	ttttactttg	ctgtgttttt	135780
gttttttctt	tttttctttt	ttattttttt	gagatggagt	ctcgctctgt	agccaggctg	135840
gactgtagtg	gcgcgatctc	ggctcactgc	aaccttggcc	tcccaggttc	aagcgattct	135900
cctgcctcag	cctcccgagt	agctgggact	ctaggcacac	accaccatgc	ccagctaatt	135960
tttgtgtttt	tagtagagac	ggggtttcac	catgttggcc	aggatggctc	caatgtcttg	136020
acctcgatg	ccgcccacct	cggctctcca	aagtgtctgg	aacacaggca	tgagccactg	136080
tgtctggcca	cattttactt	tctttgaata	tggcaggctc	acctccgtga	acaccttgag	136140
acctagtgtg	tctttgattt	taggagaagt	gggaggtgaa	tggttgagct	gtagagggtga	136200
catcagccca	gccagtggat	gggggcttgg	gaaacattgc	ttcccattat	tgtcatgctg	136260

gagggccctt	tagcccatcc	tctccccccg	ccaccctcct	tattgaggcc	tggagcagac	136320
ttcccagacc	tggtagtgt	tcagggccct	ggtagatgg	acctatat	gctgcttaag	136380
acatttgctc	ccactcaggt	tgtcccatca	gccataaggc	ccccaggag	cccgtgtgat	136440
ggagcagaga	gagacctgag	ctctgcaatc	ttgggcaagg	cttttccctt	atgtttcttc	136500
ttatctaaag	tgaacagctg	gggctcatgt	gctccctcct	catctaaagt	gaacacatgg	136560
ggctcatgtg	cagggtcctc	cccgccttca	gagcctgagg	tcccttgagg	ctcaggaagg	136620
ctgctccagg	tgagtgccga	gctgacttct	tggtggacgt	gctgtgggga	cagcccatta	136680
aagaccacat	cttggggccc	tgaattgaa	agttgtaact	gcctgggtgca	tggtggccag	136740
gcctgctgga	aacaggttgg	aagcgatctg	tcacctttca	ctttgatttc	ctgagcagct	136800
catgtgggtg	ctcactgttg	ttctaccttg	aatcttgaag	attatttttc	agaaattgat	136860
aaagttat	taaaaagcac	ggggagagaa	aaatatgccc	attctcatct	gttctggggc	136920
aggggacact	gtattctggg	gtatccagta	gggcccagag	ctgacctgcc	tccctgtccc	136980
caggctgact	gtcgaagaac	acatctggtt	ctatgcccgc	ttgaaagggc	tctctgagaa	137040
gcacgtgaag	gcggagatgg	agcagatggc	cctggatggt	ggtttgccat	caagcaagct	137100
gaaaagcaaa	acaagccagc	tgtcaggtgc	ggcccagagc	taccttccct	atccctctcc	137160
cctcctcctc	cggctacaca	catgcggagg	aaaatcagca	ctgccccagg	gtcccaggct	137220
gggtgcgggt	ggtaacagaa	acttgtccct	ggctgtgccc	ctaggctcct	tgccttcact	137280
cactgtctgg	ggctggtcct	ggagtttgtc	ttgctctgtt	ttttttagg	tggaatgcag	137340
agaaagctat	ctgtggcctt	ggcctttgtc	gggggatcta	aggttgtcat	tctggatgaa	137400
cccacagctg	gtgtggaccc	ttactccgc	aggggaatat	gggagctgct	gctgaaatac	137460
cgacaagggtg	cctgatgtgt	atttattctg	agtaaattgga	ctgagagaga	gcggggggct	137520
tttgagaagt	gtggctgtat	ctcatggcta	ggcttctgtg	aagccatggg	atactctct	137580
gttakcacag	aagagataaa	gggcattgag	actgagattc	ctgagaggag	atgctgtgtc	137640
tttattcatc	tttttgtccc	caacatgggtg	cactaaattt	atggttagtt	gaaaggggtg	137700
atgcttaaat	gaatggaagc	ggagaggggc	aggaagacga	ttgggctctc	tggttagaga	137760
tctgatgtgg	tacagtatga	ggagcacagg	caggcttggg	gccaaactctg	gcntggccct	137820
gagacattgg	gaaagtcaca	acttgccctca	ccttctttgc	cgataataat	agtgggtgct	137880
tacctcatag	aggattaaat	taaatgagaa	tgcacacaaa	ccacctagca	caatgcctgg	137940
catatagcaa	gttcccaaat	aaaatgcnta	ctgttcttac	ctctgtgagg	atgtggtacc	138000
tatatataca	aagctttgcc	attctagggr	tcatagccat	acagggtgaa	aggtggcctc	138060
caggctctctt	ccagtgtcta	cccctgctaa	tatctctcta	gtccctgtca	ctgtgacaaa	138120
tcagaactga	gaggcctcac	ctgtcccaca	tccttgtgtt	tgtgcctggc	aggccgcacc	138180
attattctct	ctacacacca	catggatgaa	gcggacgtcc	tgggggacag	gattgccatc	138240
atctcccatg	ggaagctgtg	ctgtgtgggc	tcctccctgt	ttctgaagaa	ccagctggga	138300
acaggctact	acctgacctt	ggtcaagaaa	gatgtggaat	cctccctcag	ttcctgcaga	138360
aacagtagta	gcactgtgtc	atacctgaaa	aaggtagact	gcagtcttgg	tgctgggctg	138420
gtgttggggc	tgggcagcca	ggacttgctg	gctgtgaatg	atttctccat	ctccaccctc	138480
tttgccatgt	tgaaccacc	atctccctgc	tctgttgccc	ctttgaaatc	atatcatact	138540
taaggcatgg	aaagctaagg	ggccctctgc	tcccatgtg	ctagtctgt	tgaatcccgt	138600
tttccctttc	ctatgaggca	cagagagtga	tggagaagg	ccttagagga	cattattatg	138660
tcaaagaaaa	gagacttgct	aagaggtaag	agccttggtc	ancaaagac	ctggtngttc	138720
ctgctcatta	cttttcaatc	tcattgacct	taacttttaa	actataaaac	agccaatatt	138780
tattaggcac	tgatttcatg	ccagagacac	tctgggcant	gaaagaaaag	aatgataata	138840
gttaatttta	tatagcgttg	ttaccattta	caaccttttt	ttttttttta	acctctatca	138900
tctcaattaa	agtgcagaga	gacctggga	agaaggtaac	tatatattatt	atcccagatg	138960
aggaagtga	ggcttgtagg	gaattggtag	ctgattcaag	gtcaccagc	aggtaaataa	139020
cagtgggtgg	accagaccca	attaccaggt	atgttttcct	ctgtaccgca	gtacatgcct	139080
gagatttatt	tgtgtgttga	agccagtgg	acctaattgta	tttacatccc	aacctgaaac	139140
tcctatccac	ttattttacct	tttaatgagc	ctcttaactc	aagtgcagtc	tgaggaccag	139200
cagcatcagg	atcacttggg	aacttggttag	aaattcagca	acctgggccc	agctcagacc	139260
taccgaatca	gaatctgtgc	attttaacaa	ggttcttgag	tggttgaaca	cacattaaag	139320
catgagaagc	attgaactag	acatgtagcc	aggtaaaggc	cttgccctgag	atggttgga	139380
aaggcctcat	tgcagcattc	attggcaggc	cacagttcct	ttggcagctc	tgcttcctga	139440
cctttcaccc	tcaggaagcg	aggctgttca	cacggcacac	acatgccaga	cagggtcctc	139500

tgaagccacg	gctgccagtg	catgtgtccc	agggaaagct	ttttccttta	gttctcacac	139560
aacagagctt	cttgggaagcc	ctccccggcg	aaggtgctgg	tggctctgcc	ttgctccgtc	139620
cctgacccgt	tctcacctcc	ttcttttgcca	tcaggaggac	agtgtttctc	agagcagttc	139680
tgatgctggc	ctgggcagcg	accatgagag	tgacacgctg	accatcggtg	aggactctgg	139740
ggtttcttat	tcaggtggtg	cctgagcttc	ccccagctgg	gcagagtggg	ggcagaggag	139800
gagaggtgca	ggggctgggtg	gcgctgactc	aaggtttgct	gctgggctgg	ggctgggtgg	139860
ctgcgggkgt	gggagcagct	tggtgccggg	ttggccta	gcttgctggg	gtgcctgggg	139920
ctcggtttgg	gagctagcag	ggcagtgctc	cagagagctg	agatgattgg	ggtttgggga	139980
atcccttagg	ggagtggaca	ctgaatacca	gggatgagga	gctgagggcc	aagccaggag	140040
ggtgggattt	gagcttagta	cataagaaga	gtgagagccc	aggagatgag	gaacagcctt	140100
ccagattttt	cttgggtagc	gtgtgtagga	ggccagtgtc	accagtagca	tatgtggaac	140160
agaagtcttg	acccttgcta	tctctgccta	gtccta	ctggcttttc	ccaggaaggc	140220
ttctgcttcc	atggactggt	agattaaccc	tttatttagg	taaatgaggg	aacctacttt	140280
ataagcatag	gaaaggggtg	agaatctttt	aagattcctt	tactcaagtt	ttcttttgaa	140340
gaatcccaga	gcttaggcaa	tagacaccag	actttgagcc	tcagttatcc	attcacccat	140400
ccacccaccc	accaccccat	ccttccatcc	tcccatcctc	ccattcaccc	atccacccat	140460
ccagctgtcc	accatttcta	cactgagtac	ctataatgtg	cctggctttg	gtgatacaaa	140520
ggtgaataag	acatagtcct	ttcctttgcc	cccaaccctc	agaccagaga	tgaacatgtg	140580
gaatgacctg	aacacctgga	acaggtgtgg	tgtatgagcg	gcaggcctct	gatgagaggg	140640
tgggggatgg	ccagccctca	ctccgaagcc	cctctgagtt	gattgagcca	tctttgcatt	140700
ctgggtccctg	cagatgtctc	tgctatctcc	aacctcatca	ggaagcatgt	gtctgaagcc	140760
cggctgggtg	aagacatagg	gcagtgagctg	acctatgtgc	tgccatatga	agctgctaag	140820
gagggagcct	ttgtggaact	ctttcatgag	attgatgacc	ggctctcaga	cctgggcatt	140880
tctagtattg	gcctctcaga	gacgacctg	gaagaagtaa	gttaagtggc	tgactgtcgg	140940
aatatatagc	aaggccaaat	gtcctaaggc	cgaccagta	gcctgcattg	ggagcaggat	141000
tatcatggag	ttagtctattg	agtttttagg	tcactcgacat	ctgattaatg	ttggccccag	141060
tgagccattt	aagatggtag	tgggagatag	caggaaagaa	gtgttttcct	ctgtaccaca	141120
gtacatgcct	gagatttgtg	tgttgaaacc	agtgttacct	aacacattta	catcccaacc	141180
ttaaactcct	atgcacttat	ttacccttta	atgagcctct	ttacttaagt	acagtgkgag	141240
gaacagcggc	atcaggatca	cttggggaact	tgtagaaat	tcagcaactt	gggcccagct	141300
cagacctact	gaatcagaat	caggagcaat	tctctggtgt	gactgtgtca	cagccaggta	141360
tcaactggat	tctcatacat	aggaaatgac	aaacgtttat	ggatggatag	tctacttgtg	141420
ccaggtgctg	agattttgtt	tttggttttt	gatttttttt	taatcactgt	gacctcattt	141480
aattctcaaa	aaaagatgaa	aaaatgaaca	ctcaggaatg	ctgacatgag	attcagaatc	141540
aggggttttg	ggcttcaaa	tccatcctct	ctttatccat	gtaatgcctc	cccttagaga	141600
tacaacatca	cagaccttga	aggctgaagg	ggatataaaa	gctgtctggc	caagtgggtct	141660
ccaagcttga	cagtgcagca	gaatcacctg	gggatattat	taaaaataaa	cataactaagg	141720
tttggttcca	gggcctgtga	atcagaattt	ctggaggtga	ggccttgaag	tctgtatttc	141780
tattgcatac	tttggaacaca	gtggtctata	gactagagtt	tggaaatgat	tgcgctcatt	141840
cagattctct	tctgatgttt	gaattgctgc	catcatattt	ctagtgtctc	atttcctcct	141900
gtcattctct	tcttgataaa	cttatcatag	tactagccta	ctcaaagatt	tagagccaca	141960
gtcctgaaag	aagccacttg	actcattccc	tgtaggttca	gaataaattt	cttctgcgca	142020
gtgtctgtca	tagctttttt	taaatttttt	tttatttttg	atgagactgg	agttttgctc	142080
ttattgccc	agctggagtg	cagtgggtgc	attttggtc	actgcaacct	ccacctccca	142140
ggttcaagcg	attctcctgc	ctcagcctcc	caagtagctg	agattacaag	catgtgctac	142200
cacgcccagc	taatttttga	tttttagtag	agatgggttt	tatccatgtt	ggtcaggctg	142260
gtctcgagct	ccagacctca	ggtgatctgc	ccgcctcggc	ctcccaaagt	gctgggatta	142320
taggcctgag	ccacagcgct	cagccataac	tttaatttga	aaatgattgt	ctagcttgat	142380
agctctcacc	actgaggaaa	tgttctctgg	caaaaacggc	ttctctccca	ggtaactctg	142440
agaaagtgtt	attaagaaat	gtggcttcta	ctttctctgt	cttacggggc	taacatgcca	142500
ctcagtaata	taataatcgt	ggcagtggtg	actactctcg	taatgttggt	gcttataatg	142560
ttctcatctc	tctcattttc	cagatattcc	tcaaggtggc	cgaagagagt	ggggtggatg	142620
ctgagacctc	aggtaactgc	cttgagggag	aatggcacac	ttaagatagt	gccttctgct	142680
ggctttctca	gtgcacgagt	attgttctct	tccctttgaa	ttgttctatt	gcattctcat	142740

ttgtagagtg	taggtttgtt	gcagatgggg	aaggtttgtt	ttgttgtaaa	taaaataaag	142800
tatgggattc	tttccttggtg	ccttcagatg	gtaccttgcc	agcaagacga	aacaggcggg	142860
ccttcgggga	caagcagagc	tgtcttcgcc	cgttcaactga	agatgatgct	gctgatccaa	142920
atgattctga	catagaccca	ggtctgttag	ggcaagatca	aacagtgtcc	tactgtttga	142980
atgtgaaatt	ctctctcatg	ctctcacctg	ttttctttgg	atggccttta	gccaaggtga	143040
tagatcccta	cagagtccaa	agagaagtga	ggaaatggta	aaagccactt	gttctttgca	143100
gcatcgtgca	tgtgatcaaa	cctgaaagag	cctatccata	tcacttcctt	taaagacata	143160
aagatgggtc	ctcaatcctc	tgaacccatg	tattttattat	cttttctgcg	gggtcctagt	143220
ttcttgata	cattaggtgt	ttaattgttg	aacaaatatt	cattcgagta	gatgagtgat	143280
tttgaaagag	tcagaaaggg	gaatttgctg	ttagagttaa	ttgtacccta	agacttagat	143340
atttgaggct	gggcatgggtg	gctcatgcca	gtaatcccag	cgctttgaga	ggctgagggtg	143400
ggtagatcac	ctgaggtcag	gagtttgaga	ccagtctgac	caacaagggtg	aaaccccgtc	143460
tctactaaat	acaaaaaatt	agccgagtgt	ggtggcacat	gcctgtcatc	ccagctactt	143520
gggaggctga	ggcaggagaa	tcgcttgaac	ccaggaggca	gaggttgca	tcagccacgg	143580
ttgcgccatt	gcactccaga	ctgggcaaca	agagtgaaaa	ctccatctca	aaaaagaaaa	143640
aaaaagaatt	agatattttg	gatgagtgtg	tctttgtgtg	tttaactgag	atggagagga	143700
gagctaagac	atcaaacaaa	tattgttaaag	atgtaaaagc	acatcagtta	ggtatcatta	143760
gtttaggaca	aggatttcta	gaaaattttt	aggaacagaa	aactttccag	ttctctcacc	143820
cctgctcaaa	gagtgtatgg	ctcttacatt	atatataact	gcctgacttc	atacagtatc	143880
agtacttaga	tcatttgaaa	tgtgtccacg	ttttaccaa	atataatagg	gtgagaagct	143940
gagatgctaa	ttgccattgt	gtattctcaa	atatgtcaag	ctacgtacat	ggcctgtttc	144000
atagagtagt	ctataagaaa	ttgatgactt	gattcatccg	aatggctggc	tgtaacacct	144060
ggttacgcat	gaacacctct	tttcagttgt	ctcaagacac	ctttcttttc	tgtacttatc	144120
agacaaggac	tgaaggcag	agactgctac	tgttagacat	tttgagtcaa	gcttttccct	144180
ggacatagct	ttgtcatgaa	agccctttac	ttctgagaaa	cttctagctt	cagacacatg	144240
ccttcaagat	agttgttgaa	gacaccagaa	gaaggagcat	ggcaatgccg	aaaacacctt	144300
agataatagg	tgaccttcag	tgttggtctc	ttgcagaatc	cagagagaca	gacttgctca	144360
gtgggatgga	tggcaaaggg	tcctaccagg	tgaaaggctg	gaaacttaca	cagcaacagt	144420
ttgtggccct	tttgtggaag	agactgctaa	ttgccagacg	gagtccgaaa	ggattttttg	144480
ctcagggtgag	acgtgctgtt	ttcgccagag	actctggctt	catgggtggg	ctgcaggctc	144540
tgtgaccagt	gaaggcagga	tagcatcctg	gtcaagatat	ggatgccgga	gccagattta	144600
tctgtatttc	aatcccagtt	ctattccttg	ccagttgtgt	atccgctggc	aagttacttc	144660
tctatgcctc	aatctcctca	tctgtaaaat	ggggataata	atattacctg	caatacaggg	144720
ttgttacgaa	aataaaaaatg	aatagggtgct	tagaatgggg	cctgacatta	gtaagtgtct	144780
agttttgtgt	gtgtatatgt	tatttttatt	ttggaggaga	acataaaaaag	gacaaagtgt	144840
agaaaaactg	gttgggtgta	ttcagctgtc	ataacatgag	agttgttatg	cccagatgca	144900
cttgacatgt	gaattttatta	gaaacatgat	ttttctctga	gttgatgttt	aactcaaact	144960
gatagaaaag	atagggtcaga	atatagttgg	ccaacagaga	agacttgtta	gactattgtc	145020
tgcattgtcag	tgtttgcatg	ctaacttgct	tagttagaaa	ggttaaattt	tttactctta	145080
taaaatcaag	aaatatagag	aaaagggtctg	cagagagtct	ttcatttgat	gatgtggata	145140
ttgttaagag	cgggagtttg	gagcatacag	agctcaagtt	gaatcctgac	tttgctactt	145200
attggctata	tgaccttggtg	caagctgctt	agtcctctctg	atcctcagtt	acctttgttt	145260
gttgatgatg	accattgata	acacaacccat	aaataatgac	aacatagaga	tagttctcat	145320
tatagtagtt	gttatacaga	attattcact	caatgttaat	tttctgcatt	gaaatcccag	145380
aacattagaa	ttgggggcat	tatttgaaatc	tttaaggtta	taagggaatac	atttctcagc	145440
aataaatgga	aggagttttg	ggttaactta	taaagtatac	ccaagtcatt	tttttttcag	145500
agaagatatg	gtagaaagtc	ttaggaggtt	gaagaaggaa	ttggatatatt	attctttctg	145560
agactatcat	gggagataat	gactatgggt	gtccatgatt	ggagccgttg	ctgtagagtt	145620
ggttttatta	tagtgtagga	tttgaatggg	ccatgtgttc	tcagacctca	gattaaaawg	145680
agaaaactga	ggccagtggg	gagcgtgact	tcacatgggt	acacttggtc	tagagacaga	145740
accaggattc	aggacttctg	gctcctgggtc	ctgggttcat	ggcccaatgt	agtctttctc	145800
agtcttcagg	aggaggaagg	gcaggaccca	gtgttctgag	tcaccctgaa	tgtgagcact	145860
atttacttog	tgaacttctt	ggcttagtgc	ctctgccagg	tggccataac	ctctggcctt	145920
gtgttgccag	agaaaagggtt	tagttttcag	gctccattgc	ttcccagctg	ccaagaatgc	145980

tatatnttta	tttnatztat	ttnnnnnnnn	ttttttgaga	tggagtctcg	ctctgttgcc	149280
caggctggag	tgcagtggct	caatcttggc	tcactgccac	ctctgcctcc	caggttcaag	149340
tgattctcct	acctcagcct	cctgagtagc	tgggattaca	agcatgtgcc	accacactgg	149400
ctaatttttc	taatttttagt	agagatggag	tttcaccatg	ttggacagga	tggctcttgaa	149460
ctcctgactt	agtgatccac	cgcctctcgc	ctcccaaagt	tctgggatta	caggcatgag	149520
ccatcgtgcc	tggcaattat	atttaatat	taataataag	gaaataattg	ctgtaacttt	149580
actttaaatt	gtggaattct	gaaactggaa	gggaactgga	aatgacttgt	tgaatcaa	149640
cattttaaac	ttttattttg	ccagtggaaa	aaataagccc	ccaaaagagc	aggggacctg	149700
ctgnatgtcc	cacagtaatt	cagagctgga	gatgaggttg	aaggctttgt	gtcttatctc	149760
cagggaaaat	ttgtagacag	cgtagctctn	ttatgtgacg	agcattctca	ccccagtc	149820
cccccaattc	tctactcatt	tgagaacata	aattggatct	tgccagtctc	tactcatttt	149880
tcagcacatc	gagcataaga	tccagactct	ttccagggcc	tctctcatct	ggctcctctc	149940
ctcctccttt	atcattactc	ttcttcgtag	cttatctctac	tccagccatg	ctgtcttcct	150000
attattccta	aaaartagaa	atgcatttct	tcctagggcc	tttgtacctg	cacttgccat	150060
cgcttttgct	cagaatgttc	tttttgccaa	gcttttgccc	agcttggtct	ccatcattgt	150120
tatgttttg	ctgaaatgtc	ttctcttagt	aggttcattc	tccccagtc	ctgtcttttt	150180
attttgcttt	attttgggcc	atctaagggt	atcttattag	tgtatttgtt	gttcgtctcc	150240
tccatgggca	tacacctcca	tgaaggcagg	tattttcacc	ttaggccttc	gaatatactg	150300
gacagcatct	ggcacgtagt	agatgctcaa	cgaatgtttg	ttgtgtgagc	aatgggttg	150360
ttgattgat	tgaactgagt	tcagtatgta	aatatttagg	gcctctttgc	attctatttt	150420
acttatgtat	aaaatgatac	ataatgatga	tataaatgat	gtcacagtgt	acaaggctgt	150480
tgtgggatca	agcaatcaaa	tgagatcatg	cttgtctttt	ccaaatggtg	agggaaataga	150540
tgcatgtttg	tggttgttac	ggaatgatcc	tgtgtctctg	aggcaacaga	aaggccaggc	150600
catctctggg	aatcctactc	ttgctgtctt	ccctttgcag	agacacgccc	tgccaggcag	150660
gggaggaaga	gtggaccact	gccccagttc	cccagaccat	catggacctc	ttccagaatg	150720
ggaactggac	aatgcagaac	ccttcacctg	catgccagt	tagcagcgac	aaaatcaaga	150780
agatgctgcc	tgtgtgtccc	ccaggggcag	gggggctgcc	tctccacaaa	gtgagtcact	150840
ttcagggggt	gattgggcag	aaggggtgca	ggatgggctg	gtagcttccg	cttgggaagca	150900
ggaatgagtg	agatatcatg	ttgggagggg	ctgtttcagt	cttttttgtt	ttttgttttt	150960
ttttctgagg	cggagtcttg	ctctggctgc	ccaggctgga	gtgctgtggc	atgatcttgc	151020
ctcactgcaa	cctccacctc	ccaggttcaa	gcgattctcc	tgctcagcc	tctgagtag	151080
ctgggattac	aggcacgcac	caccatgtct	ggctaatttt	tgtgttttta	gtagagatag	151140
ggtttcgccg	tgttggttag	gctggctctg	aattcctgac	ctcaggatg	cccccgcct	151200
cggcctccca	aagtgtctgg	attacaggcg	tgagccacta	cgccagccc	tgtttcagtc	151260
tttaactcgc	ttcttgtcat	aagaaaaagc	atgtgagttt	tgaggggaga	aggtttggac	151320
cacactgtgc	ccatgcctgt	cccacagcag	taaagtcaca	ggacagactg	tggcaggcct	151380
ggcttccaat	cttggctctg	caacaaatga	gctggtagcc	tttgacaggc	ctgggcctgt	151440
ttcttcacct	ctgaattagg	gaggctggac	cagaaaactc	ctgtggatct	tgtcaactct	151500
ggtattctta	gagactctgt	ttgggaagga	gtcctgagcc	attttttttt	tcttgagaat	151560
ttcaggaaga	ggagtgttta	tgatagctct	ctgctgtttt	tatcagcaac	caaattgcag	151620
gatgaggaca	agcaattcta	aatgagtaca	ggaactaaaa	gaaggcttgg	ttaccactct	151680
tgaaaataat	agctagtcca	ggtgcggggg	ggctcacacc	tgtaatctca	gtattttggg	151740
atgccgaggt	ggactgatca	cctaagggtca	ggagttcgaa	accagcttgg	ccaatgtggc	151800
gaaacctgt	ctctactaaa	aattcaaaaa	ttagccaggc	atggtggcac	atgcctgtaa	151860
tcccagttac	ttgggaggct	gaagcaggag	aattgcttga	acctgggagg	tggaggtcgc	151920
agggagccaa	aattgcgcca	ctgtactcca	gcctgagcaa	cacagcaaaa	ctccatatca	151980
aaaaataaaa	tgaataaaat	aacagcta	ctagtcata	gtataactcc	agtgaacaga	152040
agatttatta	ggcatagtga	atgatggtgc	ttcctaaaaa	tctcttgact	acaaagaatc	152100
tcatttcaat	gtttatttgt	tagatgttca	gaataaattc	ttgggaaaga	ccttggtctg	152160
gtgtaagtga	attaccagt	ccgagggcag	ggtgaaccaa	gtctcagtgc	tgggtgactg	152220
agggcagtgt	ctgggacctg	tagtcagggt	tccggtcaca	ctgtggacat	ggtcactgtt	152280
gtccttgatt	tgttttctgt	ttcaattctt	gtctataaag	acccgtatgc	ttggttttca	152340
tgtgatgaca	gagaaaacaa	aacactgcag	atatccttca	ggacctgaca	ggaagaaaca	152400
tttcggatta	tctggtgaag	acgtatgtgc	agatcatagc	caaaagggtga	ctttttacta	152460

aacttgcccc	ctgcckttatt	attactaatt	agaggaatta	aagacctaca	aataacagac	152520
tgaacacgtg	ggggaaatgc	cagattatgg	cctgattctg	tctatttgaa	gttttaggata	152580
ttatcccaaa	ctagaaaaga	tgacgagagg	gactgtgaac	attcagttgt	cagcttcaag	152640
gctgaggcag	cctggtctag	aatgaaaata	gaaatggatt	caacgtcaaa	ttttgccact	152700
tagtagcaac	ttgaccaggt	aactggttat	ccttttaaag	ccttagttta	tctaaattgt	152760
gatattaatg	ttgctcttat	aagtttgtca	tgaggactaa	attaaatggt	gtacatagag	152820
tgccttggtt	actctctgat	gggggactcc	atgataatth	gtggtctcat	ggagggagct	152880
ctgggaaggt	ttaggagcct	gccttggtct	tgcagccttg	ggagagcctt	ctagcttccc	152940
aggacatggc	agcctagtgt	tgaatgcttg	gctcagcaaa	tgtttgttct	cgtttccttc	153000
ccatcaactt	ggtcagttgg	ggtctttcag	ttaggagtat	ctcagtgact	ttaaattggca	153060
tgggcatgct	ggagtgatag	tgaccatgag	tttctaagaa	agaagcataa	tttctccata	153120
tgtcatccac	aattgaaata	ttattgttaa	ttgaaaaagc	ttctaggcca	ggcacggtgg	153180
ctcatgcctg	taatcccagc	actttaggag	gccaaggcgg	gtggatcact	tgaggtcagg	153240
agtttgagac	cagcctggcc	aacatgggga	aaccctgtct	ctactaaaaa	tacaaaataa	153300
gctgggcgtg	gtggtgctg	cctgtaatcc	cagctacttg	ggaggctgag	gcaggagaat	153360
tgcttgaatc	tgggaggcgg	aggttgcagt	gagctgagtt	catgccattg	cattccagcc	153420
tgggcaacaa	gagcgaacc	atctcccaaa	agaaaaaaa	aagaaagaaa	aagcttctag	153480
tttggttaca	tcttggtcta	taagggtggt	tgtaaattgg	tttaacccaa	ggcctggttc	153540
tcatataagt	aatagggtat	ttatgatgga	gagaaggctg	gaagaggcct	gaacacaggc	153600
ttcttttctc	tagcacaacc	ctacaaggcc	agctgattct	agggttattt	ctgtccgttc	153660
cttatatcct	caggtggata	tttactcctt	ttgcatcatt	aggaataggc	tcagtgcctt	153720
ctttgaactg	atthtttgtt	tctttgtctc	tgcagcttaa	agaacaagat	ctgggtgaat	153780
gagtttaggt	aagttgctgt	ctttctggca	cgtttagctc	agggggagga	tggtgttgta	153840
ggtgtncctg	gattgaagaa	agccttgggg	attgtttgtc	actcacacac	ttgtgggtgc	153900
catctcactg	tgaggaggac	agaagccctg	tgaacatgtg	gagcacacag	gggcacagac	153960
agatttagat	taggcctgct	ttatagagtt	tctgcctaga	gcatcatggc	tcagtgccca	154020
gcagccctc	cagaggcctc	tgaaatatth	gataactga	tttctctgag	gagaatcaga	154080
aatctcctgc	agggtgtctag	ggatttcaag	taagtagtgt	tgtgagggga	atacctactt	154140
gtactttccc	cccaaaccag	attcccagg	cttcttaagg	actcaaggac	aatttctagg	154200
catttagcac	gggactaaaa	aggctcttaga	ggaaataaga	agcgccaaaa	ccatctcttt	154260
gcactgtatt	tcaaccattt	tgtccttctg	ggttttgaag	gaacagggtg	gactggggac	154320
agaagagttc	ttgaagccag	tttgtccatc	atggaaaatg	agatagggtg	tgtggctacg	154380
tcagggggcc	cgaaggctcc	ttgttactga	tttccgtctt	ttctctctgc	cttttcccca	154440
agggccagga	cccctggatc	tctgggcaga	gcagacgcag	gcccctataa	tagccctcat	154500
gctagaargg	agccggagcc	tgtgtataag	gccagcgcag	cctactctgg	acagtgcagg	154560
gttcccactc	tcccaactcc	ccatctgctt	gcctccagac	ccacattcac	acmcgagcca	154620
ctgggttgga	ggagcatctg	tgagatgaaa	caccattctt	tcctcaatgt	ctcagctatc	154680
taactgtgtg	tgtaatcagg	ccaggctctc	cctgctgggc	agaaaccatg	ggagttaaga	154740
gattgccaac	atthattaga	ggaagctgac	gtgtaacttc	tnngaggcaa	aatttagccc	154800
tcctttgaac	aggaatttga	ctcagtgaac	cttgtaacac	ctcgactga	gtctgtgct	154860
gatgatactg	tgcacccac	tgtctgggtt	ttaatgtcag	gctgttcttt	taggtatggc	154920
ggcttttccc	tgggtgtcag	taatactcaa	gcatttctc	cgagtcaaga	agttaatgat	154980
gccatcaaac	aatgaagaa	acacctaaag	ctggccaagg	taaaatatct	atcgtaagat	155040
gtatcagaaa	aatgggcatg	tagctgctgg	gatataggag	tagttggcag	gttaaacgga	155100
tcacctggca	gctcattggt	ctgaatatgt	tggcatacag	agccgtcttt	ggcatttagc	155160
gatttgagcc	agacaaaact	gaattactta	gttgtagctt	taaaagtgtg	ggtcaaaaac	155220
aaatccagag	gccaggagct	gtggctcatg	cctgtaatcc	tagcactttg	ggaggctgaa	155280
gcggttggtg	cacttgaggt	caggagtctg	agaccagcct	ggcctacatg	acaaaacccc	155340
gtatctacta	aaaatacaaa	aaaattagct	gggcttggtg	gcacacacct	gtaatccag	155400
ctacttgagg	ggctgaggca	ggagaattgc	ttgaaccctg	taggaagagg	ttgtagttag	155460
ccaagatcgc	accgttgac	tccagcctgg	gcaacaagag	caaaactcca	tctcaaaaaa	155520
caaattaaat	ccagagattt	aaaagctctc	agaggctggg	cgcggtggct	tacacctgtt	155580
atccagcat	tttgggatgc	cgaggcgggc	aaagcacaag	gtcaggagtt	tgagaccagc	155640
ctggccaaca	tagtgaaacc	ctgtctctgc	taaaaacata	gaaaaattag	ccgggcatgg	155700

tggcgtg	ctgtaatccc	agctactcgg	gaggetgagg	tgagagaatt	rcctgaaccc	155760
gggaggcgga	ggttgacgtg	agcccagatt	gcaccactgc	actccagcct	gggcgacaga	155820
gcaagactcc	atctcaaaaa	aagctctcag	aacaaccagg	tttacaaatt	tggtcagttg	155880
gtaaataaac	tggttttcaa	acatactttg	ctgaaayaat	cactgactaa	ataggaaatg	155940
aatctttttt	tttttttttt	taagctggca	agctgggtctg	taggacctga	taagtactca	156000
cttcattttct	ctgtgtctca	ggtttcccat	ttttaggtga	gaattaaggg	gctctgataa	156060
aacagaccct	aggattgtgg	acagcagtr	tagtcctaga	gtccacaagt	ctgcttttga	156120
gtgatgggccc	catgtatctg	gcacatctgc	aggcagagcg	tggttctggc	tcttcagatg	156180
atgccgggtgg	agcactttga	ggagtccctca	ccccaccgtg	ataaccagac	attaaaatct	156240
tggggctttg	catcccagga	tttctctgtg	attccttcta	gacttgtggc	atcatggcag	156300
catcactgct	gtagattttct	agtcacttgg	ttctcaggag	ccgtttatatt	aatggcttca	156360
catttaattt	cagtgaacaa	ggtagtggca	ttgctcttca	cagggccgctc	ctgttgtcca	156420
caggttccag	attgactggt	gccccttatc	tatgtgaaca	gtcacaactg	aggcagggtt	156480
ctgttgttta	caggacagtt	ctgcagatcg	atctctcaac	agcttgggaa	gatttatgac	156540
aggactggac	accaraaata	atgtcaaggt	aaaccgctgt	ctttgttcta	gtagcttttt	156600
gatgaacaat	aatccttatg	tttctggag	tactttcaac	tcatggtaaa	gttggcaggg	156660
gcattcacaa	cagaaaagag	caaactatta	actttaccag	tgaggcagta	cgggtgtagtg	156720
tagtgattca	gagaatttgc	tttgccacca	gacataccag	gtaaccttga	ctaagttact	156780
taacctatct	aaacctcagt	tycctcatct	gtgaaatgga	gacagtaatc	atagctatatt	156840
ccaaactggt	gtgagaattc	aatgagttaa	aggtataagg	tcctcaccac	agcgccgtgc	156900
cacatagtca	gtgatcacta	tgtcctgaac	actgtaatta	cttcgccata	ttctctgac	156960
atagtgtttt	gccttgggtat	gtgactagaa	tttctttctg	aggtttatgg	gcatggttgg	157020
tgggtatgca	cctgcctgca	ggagcccggc	ttgggggcat	tacctgtac	ctgggtatgtt	157080
ttcttttcagg	tgtggttcaa	taacaagggc	tggcatgcaa	tcagctcttt	cctgaatgtc	157140
atcaacaatg	ccattctccg	ggccaacctg	caaaaggag	agaaccttag	ccattatgga	157200
attactgctt	tcaatcatcc	cctgaatctc	accaagcagc	agctctcaga	ggtggctctg	157260
taagtgtggc	tgtgtctgta	tagatggagt	ggggcaagg	agagggttat	ggagaagggg	157320
agaaaaatgt	gaatctcatt	gtaggggaac	agctgcagag	accgttatat	tatgataaat	157380
ctggattgat	ccaggctctg	ggcagaagtg	ataagtttac	gaattggctg	gttgggcttc	157440
ttgaactgca	gaagagaaaa	tgacactgat	atgtaaaaat	cgtaacattt	agtgaattca	157500
tataaagtga	gttcaaaaat	tgttaattaa	attataattt	aattataagt	gtttaatcag	157560
tttgatttgt	ttaaaaacca	ctgtttttaa	tttggtggaa	tatgttttta	ttagcttgta	157620
tctttaattc	ctaaattaag	ctgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	157680
gtgtgtgnnt	gtgaagttta	aagccaggat	gagctagttt	aaagtatgca	gcctttggag	157740
tcatacagat	ctgggtttga	atctggtctc	taaactttat	agatgtatga	tattaaatga	157800
ggcagttcat	gtaaattgcc	aagcccagca	ctcagcacag	agttgatatt	tcacacacat	157860
tagatacctt	tcctgtatgt	ggagcatggc	agttcctgtt	tctgctttac	tcctacagga	157920
tactaatata	ggacactagg	atctttatac	caagacccca	tgtaatgggc	ttatgagacc	157980
attctttctta	taaaaatctg	acagaatttt	tgtatgtgtt	agatcaatag	gctgcatact	158040
gttatttttca	agttgattta	cagccagaaa	tattaattta	tttgagtagt	tacagagtaa	158100
tattttctgct	ctcatttagt	tttcaagccc	cactagtcct	ttgtgtgtga	aaatttacia	158160
cttactgctc	ttacaaggtc	atgaacagtg	gaccaaagtg	aatgccatta	accactctga	158220
cttctctcat	tagttttatt	gtgacagtgg	actcttttga	cctcagtaat	accagtttgg	158280
cattttacatt	gtcatatttt	tagactttaa	aatgatcatc	tttaacctga	ataaaatgtg	158340
tctggtgaac	agatgttttt	ccttgggctg	tgcctcagat	atctctgtgt	gtgtgtacgt	158400
gtgtgtttgt	ctgtgtgtcc	atgtcctcac	tgattgagcc	ctaactgcat	caaagacccc	158460
tcagattttc	acacgctttt	tctctccagg	atgaccacat	cagtggatgt	ccttgtgtcc	158520
atctgtgtca	tottttgcaat	gtccttcgtc	ccagccagct	ttgtcgtatt	cctgatccag	158580
gagcgggtca	gcaaagcaaa	acacctgcag	ttcatcagtg	gagtgaagcc	tgtcatctac	158640
tggtctctca	atthttgtctg	ggatatggta	aggacacagg	cctgctgtat	ctttctgatg	158700
tctgtcaggg	ccatggattg	atatggataa	gaaagaaaga	gctctggcta	tcatacaggaa	158760
atgttccagc	tactctaaag	atgtatgaaa	aagaaatagc	cagaggcagg	tgatcacttt	158820
catgacacca	aacacagcat	tgggtaccag	agttcatgtc	acaccagagg	gaaaattctg	158880
tacacaatga	tgaaaattaa	taccactacc	acttaagttc	ctatgtgaca	actttcccaa	158940

gaatcagaga	gatacaagtc	aaaactccaa	gtcaatgcct	ctaactttctc	tgatggggttt	159000
taacctccag	agtcagaatg	ttctttgcct	tactaggaaa	gccatctgtc	attnngaaaa	159060
ctctgtacat	tttatcagca	gcttatccat	ccattgcaaa	tantgttttt	gtgccagcca	159120
caatatattg	cttctatttg	gaccaatatg	ggggatttga	aggaattctg	aagttctaata	159180
tatatattcaa	ctctacttta	caatatctcc	ctgaaatata	tctccctgta	acttctatta	159240
attataagct	acacagagca	aatctaattc	ttctcccacc	gaacaagtcc	ctggatattt	159300
aaaaataact	ctcatactct	cattttaacct	gagtattacc	cagataagat	gatatatgag	159360
aatacacctt	gtaacctccg	aagcactgta	caaagtgtgag	caatgatggg	ggagatgatg	159420
atgagatctt	tgctgtttat	accaagcccc	ttagactgtg	tcactcttct	gatccgggtg	159480
tccttgtagt	gccatgctgt	atattgtgaa	tgtcccgttt	tcaaaagcaa	agccaagaat	159540
taaccttgtg	ttcaggctgt	ggtctgaatg	gttatgggtc	cagagggagt	tgatcttttag	159600
ctcacacttc	tattactgca	gcacaaagat	tttgcathtt	ggaaggagca	ccgtcttact	159660
ggcaacttag	tggtaaacca	aaacctccat	ttcacacaaa	tgattgtgaa	attcgggtct	159720
ccttcattct	atacaaattc	atttgatttt	tttgaaacta	aactttatat	ttatccatat	159780
taaattacat	gggtttttatt	tttgttttat	cttgattcag	taattactcc	tttcagtaaa	159840
cacagactga	gtgctgtgtg	tctgacttat	gccaggcata	ggtgattcag	agatgaaagg	159900
tcaagtcctt	gaacctcatct	cttgtcttcc	tgggtattat	ctgtccctcc	ctgcttttaga	159960
gctcctgaaa	tttgctagaa	gcattgtctt	atctaagttg	ttgataaaca	catcaagtag	160020
gattggactg	aggcagagcc	ctgtagtctg	aagctgcagt	tcttctagcg	gctgacaagc	160080
cccactatca	cttccctgct	ggtgctttgc	tctgccagct	gtgaattctc	ataattgtcc	160140
tatcgtcaag	tctttatttc	tgcattttac	tgcttgatac	actgtcagga	cagactttaa	160200
aattattctc	agtgcgatga	aacaattctg	acattcatgt	tatgagcagt	tacctcataa	160260
atagattaca	tgtgagattg	aacttgggca	gactataata	tagcattaat	gatgaaacag	160320
acacagtcac	cttcgggaag	aagaatagag	gcttatttgc	tgcctgtgaa	attaaaatta	160380
ctctgactgg	gaatccatcg	ttcagtaagt	ttactgagtg	tgacaccttg	gcttgactgt	160440
tggaaaagaca	gaaagggcat	gtagttttata	aaatcagcca	aggggaaaat	gcttgtcaaa	160500
atgtattgtc	gggtattttg	attaatagtt	tatgtggctt	cattaattca	gagttactct	160560
ccaatatggt	tatctgccct	ttcttgtctg	ataatgggtg	aaacttgtgt	gatgcattgt	160620
atatttgatt	taggggtgaa	ctggatgtct	ttgttttcac	ttttagtgca	attacgttgt	160680
ccctgccaca	ctgggtcatta	tcactctcat	ctgcttccag	cagaagtcct	atgtgtcctc	160740
caccaatctg	cctgtgctag	cccttctact	tttgctgtat	gggtaagtca	cctctgagtg	160800
agggagctgc	acagtggata	aggcatttgg	tgcccagtg	cagaaggagg	gcagggactc	160860
tcagtagaca	cttatctttt	tgtgtctcaa	caggtgggtca	atcacacctc	tcattgtaccc	160920
agcctccttt	gtgttcaaga	tccccagcac	agcctatgtg	gtgctcacca	gcgtgaacct	160980
cttcattggc	attaatggca	gcgtggccac	ctttgtgctg	gagctgttca	ccgacaatgt	161040
gagtcattgca	gagagaacac	tcctgctggg	atgagcatct	ctgggagcca	gaggacagtg	161100
tttaattgtg	atcttattcc	acttgtcagt	ggtattgaca	ctgctgactg	ccttgtcctg	161160
tcttcagagt	ctgtcttccc	tgagaaggca	aagcaccttt	ctttcttgc	gtgccttaca	161220
ttttgctgg	caagcctttc	agtttctttt	gacagttttt	tttacttctt	tcttttttca	161280
atgttgctct	taccaagagt	agctcctctg	ccttccactt	tacacatgag	agctgggcga	161340
cgncattcag	tcctaaggct	tttaccatca	cctctcttgg	tgtttttatt	gtcatctcta	161400
agatcaatgc	cttttagcctt	gatcataacc	ttgaactcta	atctcaaatt	ctcacttgcc	161460
tagtggattg	ctccatttag	atagtatata	gataccccaa	cctggatatg	tcctagtttt	161520
ctttccctct	ggaacttaat	gcttttcttg	ccatccctgt	cacactcagt	ggcactacca	161580
tcactcggg	tgcccaagct	ggctctttaga	gttatcctag	atgcttgctt	tgctgttgca	161640
gatttcccac	attcaactgg	ttatgttgtc	agttcttcca	ggtatggacc	tctaaaataa	161700
ggcttccct	ccattccggg	tgtcattgcc	tttgtccaaa	cacagcacac	aaggcctttt	161760
acagttgcac	aactcttctc	gtccataccc	accacacctt	ttcccagctg	taagcttcag	161820
atgagttgcc	tccaaccacc	atgctcctgt	aggcctggct	tgaaatgcc	ttcttctgtc	161880
acagggctctg	gtagtatatc	ccttgccctt	caagatttag	ctaaaatgtg	aagctttcct	161940
tacctgctgg	gaggtgttct	ctcttttctc	tgtgctctca	gagtccttag	tccatgcctc	162000
cagtacaacg	tacatccact	tacatggtaa	tttctgtgtt	acatactttt	cctactcgga	162060
gtggagctctg	tttcttaata	attttgcctc	tcccatgcc	tagcacagtg	catccagcgt	162120
atagcccctt	attcagttgg	tagatatattg	gccactgttg	ccttgtggga	tcataagttc	162180

tgatgtat	gagaagaatt	tctaaaattc	tgacaaaatc	ctgaaactca	aatattgacc	162240
cagacatgag	caatttgctt	ttcaaatgct	aagggat	taatggattt	gctttaatta	162300
aatctagcct	gtttctaagc	tttattcatt	atctctccat	actcagagca	tttctccaga	162360
ttttctaaag	aatagaattt	tattgctaca	tatcatcagc	tatgcctgct	gctattttaat	162420
tggtatctga	attaaaaggt	ctggtttgct	cctagagaa	caaatttttt	cttctactccc	162480
atatttcaga	acttgataca	tttttaggat	aaaccatgaa	tgacacccgt	ttcttctccc	162540
tcaccctccc	ttccctccca	tttttttttt	tttttttttt	tagaagctga	ataatatcaa	162600
tgatatacctg	aagtccgtgt	tcttgatcct	cccacatttt	tgcttgggac	gagggctcat	162660
cgacatggtg	aaaaaccagg	caatggctga	tgccctggaa	agggttggtg	agtgaagcag	162720
tggtcttagg	atgctttaat	ggagatggca	ctctgcatag	gccttgggtac	cctgaacttt	162780
gttttgga	gaagcaggtg	actaagcaca	ggatgttccc	ccaccccoat	gccagtgac	162840
agggctcatg	ccaacacagc	tggttggtggc	atgggttttg	tgacacaacc	atgtgtctgt	162900
gtctctgata	gcattgagaa	aagtgaagg	gcagttttga	aggtaaggaa	aatagtgtta	162960
tttgcttgga	tccactggct	catgccactg	tctgggttg	ttagaagcac	tggaaaagtc	163020
aaaccataac	tttgagaatt	agggtgatcag	ggaatcagaa	ggaaagatgc	aaactttggc	163080
tcttttaggc	gaatcatgtg	cctgcagatg	aggctattta	ttatctttta	cacagtctat	163140
aaaattataa	tgtattacat	ctttttctac	ctttagaatg	gttaaaaata	tttctccggt	163200
agccatatga	ttattattca	tccattagat	aatatagtca	aatgggccat	gttattttact	163260
gttcatagaa	gaggggcttt	ttgcaacttg	ggctacaaag	gagatatgta	aggaatttaa	163320
ggaatggtta	catggaacta	gatttaattg	aatctagtgg	tttaattgat	tcactaggat	163380
atatgtctact	gaaaggggaa	tctgcttaaa	gtgctttctg	atatttatta	ttactaaaac	163440
ttagaatttta	ttaaaaatac	tgactgtgaa	aattacttgg	gtcgtttgcc	tttttaaaag	163500
gatttttggc	atgtctcatt	aaaaaaagaa	atactagata	tcttcagtga	agttacaaat	163560
cgaatacaca	ttggctctga	aattctgatt	gatactgggt	cataaaaagt	tttcccaa	163620
cagacttggga	aagtgatcac	tctcttggtta	ctcttttttc	cttgtcatgg	gtgatagcca	163680
tttgtgttta	ttggaagatc	ggtgaatttt	aaggaacata	ggcccaaatt	tgaggaagg	163740
ccatgggttt	tgatccctcc	attctgaccg	gatctctgca	tttgtgtctac	taggggagaa	163800
tcgctttgtg	tcaccattat	cttgggactt	ggtgggacga	aacctcttcg	ccatggccgt	163860
ggaaggggtg	gtgttcttcc	tcattactgt	tctgatccag	tacagattct	tcacaggcc	163920
caggtgagct	ttttcttaga	accctgggag	cacctggttg	agggtcacag	aggaggcgca	163980
cagggaaaca	ctcaccaatg	ggggttgcat	tgaactgaac	tcaaaatatg	tgataaaact	164040
gattttcctg	atgtgggcat	cccgcagccc	cctccctgcc	catcctggag	actgtggcaa	164100
gtaggtttta	taatactacg	ttagagactg	aatctttgtc	ctgaaaaata	gtttgaaagg	164160
ttcatttttc	ttgttttttc	ccccaaagacc	tgtaaatgca	aagctatctc	ctctgaatga	164220
tgaagatgaa	gatgtgaggc	gggaaagaca	gagaattcct	gatggtggag	gccagaatga	164280
catcttagaa	atcaaggagt	tgacgaagg	gagagagtac	aggttacaat	agctcatctt	164340
cagttttttt	cagctttatg	tgctgtaacc	cagcagtttg	ctgacttgct	taataaaagg	164400
gcattgtgttc	ccaaaatgta	catctatacc	aaggttctgt	caattttatt	ttaaaaacac	164460
catggagact	tcttaaagaa	ttcttactga	gaattctttt	gtgatatgaa	ttcccatctt	164520
cgaatacatt	ggttttata	gcttacattt	atgtgttagt	tattaaaaca	tactaatatt	164580
gtatatctag	tcaaaactga	ggtagagaga	ataaatgggt	gattttgagt	ttgagtttca	164640
tagtccaaaa	agctgatata	ttgcctgtgt	tcaagagggt	ctatatcagc	cctctagatg	164700
ccagcatctc	caaattttac	tttttggaa	tctgtacagt	atgttgcaata	tttttattac	164760
aaatttctac	tctgtggaat	ttaattttta	aaatacctgc	aatacatata	tatgttgaat	164820
agatgaaaaa	ttatgtagat	rataatgaat	gatacggttc	taaaaagaca	ggttaaaaag	164880
taagttcact	tttattttga	gcttcagaat	cattcagaag	ccagtcgcca	caaacgcaga	164940
ccaaggctct	tggcacatca	aatatgccta	tggcttaggg	ttattgacaa	gtcttatgtt	165000
gcagtgtatg	tggtttatag	tctgccttc	cacagttgct	tgggagagct	gtgagtcact	165060
gaggcttatg	aatgtttaca	ttttgtttgt	tgcatagata	tagaagggaag	cggaagcctg	165120
ctgttgacag	gatttgctgt	ggcattcctc	ctgggtgagg	aaagacactt	tgtctatatt	165180
gcgtttgtcc	ctattagttc	agactatctc	tacccaatca	agcaacgatg	ctcgtttaaga	165240
ggtaaaagtg	gatttttaag	gcttctgtat	ttatgccagg	atggagcaat	tagtcatcga	165300
gaagagaggg	accctgtatg	tcaagagaat	gatttcagag	aatccaatac	aatttaagaa	165360
aaagcatggg	gctgggcgca	gtgattcact	cctgtaatcc	cagcactttg	ggaggccgag	165420

gtgggaggac	tcacgaggtc	aggagattga	gaccatcctg	gccaacatgg	tgaaacccca	165480
tctctactat	aaatacaaaa	attagctggg	catagtagtg	cattcctgta	gtcccagcta	165540
ctcgggaggg	tgaggcagga	gaattgcttg	aacctaggag	ggggagggtg	cccagattgc	165600
gctgctgcac	tccagcctgg	tgacagagtg	agactcatgt	caacaacaaa	aacagaaaaa	165660
gcacgcacat	ctaaaacatg	cttttgtgat	ccatttggga	tggtgatgac	attcaaatag	165720
ttttttaaaa	atagattttc	tcctttctgg	tttccgtttg	tgttctttta	tgcccttttg	165780
ccagagtagg	tggtgcnaat	ttggctange	tggttttcat	tactgttttt	cacnacnatt	165840
aacntttggc	ctcaacttga	caactcaaat	aatatattata	aatacagcca	cacttaaaat	165900
ggtcccat	tgaaatacat	atttaaata	ctatacga	tgttaaaacc	aagaaaatat	165960
ttgattcttc	tctgatattt	aagaattgaa	ggtttgaggt	agttacgtgt	taggggcatt	166020
tatattcatg	tttttagagt	ttgcttatac	aacttaatct	ttccttttca	gtgctttggg	166080
ctcctgggag	ttaatggggc	tggaataatca	tcaactttca	agatgttaac	aggagatacc	166140
actgttacca	gaggagatgc	tttccttaac	aaaaataggt	gagaaaagaa	gtggcttgta	166200
ttttgctgca	aagactttgt	ttttaattta	tttaaagaaa	taggttggtta	tttttgatta	166260
cagtggatt	tttagagttc	ataaaaaatgt	tgaaatatag	taaagggtaa	agaagcacat	166320
aaaatcatcc	atgatttcaa	tatctagaga	taatcacaa	ttacatttcc	tttcagtctc	166380
attctcttct	tttaacagct	ttattcaggt	ataatttaca	tacaataata	tttgcttggt	166440
ttttaagagt	ataatttagt	gatttttggg	aaattgagag	ttttgcaacc	atcaccacaa	166500
tccagtttta	gaacttttcc	atcacccccac	atctgtctta	tatacacata	taaatgtgcc	166560
atacaattga	gatcatactg	tatgtagaat	ttaaaattag	tttttattgt	taatgagtgt	166620
attatgaata	tttcccagtg	ggttacattt	cctaagatgt	ggaattttac	attgctacat	166680
aaaatcccc	tatgtacatg	tacctataat	ttatttaata	aattccttat	aaatggtgga	166740
cacattagtt	tccatttttc	actatgtaaa	tatgtccctg	tatacatctt	ttattatttc	166800
ctcaggaaca	attcctacaa	agtaaattgc	cctctctaaa	gagcatacaa	attgactgag	166860
ccaccgttag	gccattttct	gagactgcac	aggtcacaaa	gcaatctgat	ctttgggaat	166920
acagctacat	tttataggct	tcttagataa	tgttactcta	agtactttta	atatgtgggg	166980
cttctctggg	cttttttttt	tttgagacgg	agtttcactc	ttactgcca	ggctggagag	167040
caatggcgcg	accttggtct	actgcaacct	ccgcctccca	ggttcaagcg	attctcctgc	167100
ctcagcctcc	tgagtagctg	agattacagg	tgcccgccac	aatgcctgcc	taattttttt	167160
gtattttcag	tagagatggg	gtttcaccat	gttggccaga	ctggtctcga	gctcctgacc	167220
tcaggtgatc	cacctgcctc	agcctcccaa	agttctggga	ttacaggcat	gagccactgc	167280
gcccggcttc	tctggactta	ttatgtggag	agatagtaca	aggcagtggc	tttcagagtt	167340
ttttgaccat	gaccgttggtg	ggaaatacat	tttatatctc	aacctagtat	gtacacacag	167400
acatgtagac	acatgtataa	cctaaagttt	cataaagcag	tacctactgt	tactaattgt	167460
agtgcactct	gctattttct	attctacctt	atactgcgtc	attaaaaaag	tgctggtcat	167520
gaccactaa	atatttttcc	caaaccacta	atgaacaatg	actcacaatt	tgaacacact	167580
ggacaggggg	atagccaata	aaattgaaaa	gagcaaggaa	attaatgtat	tcatgatctc	167640
ctctcctgtc	tcttacattt	ttgcagtagc	aatgtaaagg	aatcctaaga	gaacagacat	167700
tctgggaata	gcaggcctag	cgctgcacaa	ctgctttcct	aggcttgctc	ctagtaccaa	167760
gctcctgacg	catatagcag	tggcagtaat	aaccagccca	tagtaagggt	tgtcacaggg	167820
actggttgta	agaactgatt	tggttggtat	agctgtgagg	gcctggcacg	gtgtccacgt	167880
gtgcctcaat	cctaattctg	aaaaaggctg	accctggggg	tgctaattag	atacacagag	167940
aggaatgaat	gctgccagaa	ggccaagttc	atggcaatgc	cgctgtggct	gaggtgcagt	168000
catcagtctg	gaacgtgaac	actgaacttc	tctcacatgt	gattcttcac	ttgactggct	168060
tcatagaacc	ccaaagccac	cccaccacca	cataaattgt	gtctctaggt	tctgtgttgc	168120
tcacactcaa	aatttctggg	ccttctcatt	tggtgcatgt	gaatggtgca	tatgagtga	168180
gtctaggatg	gggccttagc	gttaaagccc	tggggtagtg	tgactgagat	tggtggtaaa	168240
gaatgtgcag	tggttggtcat	gacctcagaa	attctgaaat	gggactgcac	ctgcagactg	168300
aagtgttcag	agagccaggg	aggtgcaagg	actggggagg	gtagaggcag	gaaccctgcc	168360
tgccaggaag	agctagcatc	ctgggggcag	aaaggctgtg	ctttcaagta	gcagcagatg	168420
tattggtatc	tttgtaatgg	agaagcatac	tttacaggaa	cattaggcca	gattgtctaa	168480
ccagagtatc	tctacctgct	taaaatctaa	gtagttttct	tgctccttgc	agtatcttat	168540
caaacatcca	tgaagtacat	cagaacatgg	gctactgccc	tcagtttgat	gccatcacag	168600
agctgttgac	tgggagagaa	cacgtggagt	tctttgccc	tttgagagga	gtcccagaga	168660

aagaagttgg	caaggtactg	tgggcacctg	aaagccagcc	tgtctccttt	ggcatcctga	168720
caatatatac	cttatggctt	ttccacacgc	attgacttca	ggctgttttt	cctcatgaat	168780
gcagcagcac	aaaatgctgg	ttctttgtat	ctgctttcag	gggtggaaacc	tgtaacggtg	168840
gtggggcagg	gctgggtggg	cagagagggg	gtgctgctcc	caccacacga	gtcccttctc	168900
cctgctttgg	ctcctcacca	gttgtcaggt	tatgattata	gaatctagtc	ctactcagtg	168960
aaagaacttt	catacatgta	tgtgtaggac	agcatgataa	aattcccaag	ccagaccaa	169020
gtcaaggtgc	tttttatcac	tgtaggttgg	tgagtgggcg	attcggaaac	tgggcctcgt	169080
gaagtatgga	gaaaaatatg	ctggtaacta	tagtggaggc	aacaaacgca	agctctctac	169140
agccatggct	ttgatcggcg	ggcctcctgt	ggtgtttctg	gtgagtataa	ctgtggatgg	169200
aaaactgttg	ttctggcctg	agtggaaaac	atgactgttc	aaaagtccca	tatgtccagg	169260
gctgttgat	gattggcttg	tcttccccca	gggacagcag	agcaaccttg	gaaaagcaga	169320
gggaagcttc	tcccttggca	cacactgggg	tggctgtacc	atgcctgcag	atgctcccaa	169380
atagaggcac	tccaagcact	ttgtttctta	gcgtgattga	ggctggatat	gtgattttgat	169440
ctttctctgg	aacattcttt	ctaatacatc	ttgtgttcac	tccctgaaaa	tgaagagtgt	169500
ggacacagct	ttaaaatccc	caaggtagca	actaggtcac	agttccnta	cacacggata	169560
gatgaaaaac	agatcagact	gggaagtggc	ccttgacctt	ttttcttctg	tagataagag	169620
cattgatgtt	attacgggaa	gaagcctttg	aggcttttat	gtattccacc	tccgtctgga	169680
atgtgtttct	gtaaggctaa	cagttgcaat	atactagggt	aatctgagtg	agctggaatt	169740
aaaaaaaaaa	aggaatttca	ccccaatctt	atactgactt	caatagaggt	ttcagacaaa	169800
aagttgtttt	gtatataact	atcagtcacg	aaaagataat	tacaactaaa	tggccttttt	169860
ccttccctat	ttatttggag	aaatttaatt	acataaaaaa	gtactcagaa	tatttgagtt	169920
tcttgcacac	ataagacatt	tataataatg	accttgttta	caaatgaatt	tgaaagttac	169980
tctaattctt	tgattcatca	agaaataact	agaatggcaa	gttaaaattt	aagctgtttc	170040
aaagatgctt	ctgcatttaa	aaacaaattt	atctttgatt	ttttttcccc	ccagcaataa	170100
agacttattt	tattctaatt	acaggatgaa	cccaccacag	gcatggatcc	caaagcccgg	170160
cggttcttgt	ggaattgtgc	cctaagtgtt	gtcaaggagg	ggagatcagt	agtgtttaca	170220
tctcataggt	ccgtagtaaa	gtccttgggt	cctcactgtg	ggatgtttta	actttccaag	170280
tagaatatgc	gatcattttg	taaaaattag	aaaatacaga	aaagcaaaga	gtaaaacaat	170340
tattacctga	aattatatac	gcataattct	acaaaaatgc	aagcccagta	tataactatc	170400
tctttttcac	ttaatatatt	gtaaacatta	ttccaagtca	gtgcatttag	gtgtcatttc	170460
ttatagctgg	atagtattcc	attaggatat	actcttattt	aactattccc	cctttttag	170520
acatttggat	tatttccaac	ttgttcacaa	ttgtaaacac	cactacactg	aacagcatca	170580
tccctatata	cacatgtact	tgtaacagaa	tacaattccc	taggaagctg	gaatgctgga	170640
agtcattggt	atgttctcat	ggttacagag	aatctctcta	aaactaaaac	ctctttctgt	170700
tttaccgcag	tatggaagaa	tgtgaagctc	tttgacttag	gatggcaatc	atggtcaatg	170760
gaagggttcag	gtgccttggc	agtgtccagc	atctaaaaaa	taggtaataa	agataatttc	170820
tttgggatag	tgccatagta	gaaggcttga	tatttattct	tttgtgagta	tataaatggt	170880
gcctctaaaa	taaagggaaa	taaaactgag	caaaacagta	tagtggaaaag	aatgagggct	170940
ttgaagtccg	aactgcattc	aaattctgtc	tttaccattt	actggttctg	tgactcttgg	171000
gcaagttact	taactactgt	aagagttagt	ttccctggaa	gatctacctc	ctagctttgt	171060
gctatagatg	aaatgaaaaa	aattttacatg	tgccagtact	ggtgagagcg	caagctttgg	171120
agtcaaacac	aaatgggttt	gcacccctggc	cctaccaatt	atgagctctg	agccatgggc	171180
aagtgaacta	ctccctgggc	ctcagtttct	tgtaaacatc	tgtcagactt	catgggtcca	171240
ggtgaggatt	aaaggagatc	atgtattttac	agcacatggc	atggtgcttc	acataaaaata	171300
agtatttagt	aaatgataac	tggttccttc	tctcagaaac	ttatttctgg	gcctgccagg	171360
ggccgcctct	tttcatggca	caagttgggt	tcccagggtt	cagtattctt	ttaaatagtt	171420
ttctgggatg	cctccatttg	ggtatttttt	cctgctttca	ggtttggaga	tggttatata	171480
atagttgtac	gaatagcagg	gtccaacccg	gacctgaagc	ctgtccagga	tttctttgga	171540
cttgcatctc	ctggaagtgt	tctaaaagag	aaacaccgga	acatgctaca	ataccagctt	171600
ccatcttcat	tatcttctct	ggccaggata	ttcagcatcc	tctcccagag	caaaaagcga	171660
ctccacatag	aagactactc	tgtttctcag	acaacacttg	accaagtaag	ctttgagtgt	171720
caaaacagat	ttactttctc	gggtgtggat	tctgcccccg	acactcccgc	ccataggtcc	171780
aagagcagtt	tgtatcttga	attgggtgctt	gaattcctga	tctactattc	ctagctatgc	171840
tttttactaa	acctctctga	acctgaaaaa	ggagatgatg	cctatgtact	ctataggatt	171900

attgtgagaa	tttactgtaa	taataacat	aaaaactacc	atttagtgag	cacctacat	171960
gggccaggca	ttttacttgg	tgccataatcc	tattttaaatt	agataaaaaa	gtaccaata	172020
ggtcctgaca	cttaagaagt	actcagtaaa	tattttcttcc	cctcttccct	ttaatcaaga	172080
ccgtatgtgc	caaagtaaat	ggatgactga	gcagttgggtg	atgtaggggt	ggggggcgat	172140
atagaaagtc	agtttttggc	cgggcggtgt	ggctcatgcc	tgtaatccca	gcactttggg	172200
aggctgagga	gcaggcagat	catgaggtca	ggagatccag	ataatcctgg	ccaacagggt	172260
gaaaccccg	ctctactaaa	aatacaaaaa	ttagctgggc	atgggtggtgc	gcacttgtag	172320
tcccagctac	ttgcgaggct	gaggcaggag	aattgctcga	acccaggagg	tggaggttac	172380
agtgagccaa	ggtctcgcca	ctgcaactcca	gcctggggac	agagcaagac	cccatttcaa	172440
gggggggaaa	aaagtctatt	tttaagttgt	tattgctttt	ttcaagtatt	cttccctcct	172500
tcacacacag	ttttctagtt	aatccattta	tgtaattctg	tatgctccta	cttgacctaa	172560
tttcaacatc	tggaaaaata	gaactagaat	aaagaatgag	caagttgagt	ggtatttata	172620
aaggtccatc	ttaatctttt	aacaggtatt	tgtgaacttt	gccaggacc	aaagtgatga	172680
tgaccactta	aaagacctct	cattacacaa	aaaccagaca	gtagtggacg	ttgcagttct	172740
cacatctttt	ctacaggatg	agaaaagtga	agaaaagctat	gtatgaagaa	tcctgttcat	172800
acgggggtggc	tgaaagttaa	gaggaactag	actttccctt	gcaccatgtg	aagtgttgtg	172860
gagaaaagag	ccagaagtgt	atgtgggaag	aagtaaactg	gatactgtac	tgatactatt	172920
caatgcaatg	caattcaatg	caatgaaaac	aaaattccat	tacaggggca	gtgcctttgt	172980
agcctatgtc	ttgtatggct	ctcaagtga	agacttgaat	ttagtttttt	acctatact	173040
atgtgaaact	ctattatgga	acccaatgga	catatgggtt	tgaactcaca	cttttttttt	173100
tttttttgtt	cctgtgtatt	ctcattgggg	ttgcaacaat	aattcatcaa	gtaatcatgg	173160
ccagcgatta	ttgatcaaaa	tcaaaaaggt	atgcacatcc	tcattcacta	agccatgcca	173220
tgcccaggag	actggtttcc	cggtgacaca	tccattgctg	gcaatgagtg	tgccagagtt	173280
attagtgcga	agtttttcag	aaagtttgaa	gcaccatggg	gtgtcatgct	cacttttgtg	173340
aaagctgctc	tgctcagagt	ctatcaacat	tgaatatcag	ttgacagaat	ggtgccatgc	173400
gtggctaaca	tcctgctttg	attccctctg	ataagctgtt	ctgggtggcag	taacatgcaa	173460
caaaaatgtg	ggtgtctcca	ggcacgggaa	acttggttcc	attgttatat	tgctctatgc	173520
ttcgagccat	gggtctacag	ggtcatcctt	atgagactct	taaatatact	tagatcctgg	173580
taagaggcaa	agaatcaaca	gccaaactgc	tggggctgca	agctgctgaa	gccagggcac	173640
gggattaaag	agattgtgcg	ttcaaaccct	gggaagcctg	tgcccatttg	tcctgactgt	173700
ctgctaacat	ggtacactgc	atctcaagat	gtttatctga	cacaagtgtg	ttatttctgg	173760
ctttttgaat	taatctagaa	aatgaaaaga	tggagtgtga	ttttgacaaa	aatgtttgta	173820
cttttttaatg	ttatttggaa	ttttaagttc	tatcagtgc	ttctgaatcc	ttagaatggc	173880
ctctttgtag	aaccctgtgg	tatagaggag	tatggccact	gccccactat	ttttattttc	173940
ttatgtaagt	ttgcatatca	gtcatgacta	gtgcctagaa	agcaatgtga	tggtcaggat	174000
ctcatgacan	ttatatttga	gtttctttca	gatcatttag	gatactctta	atctcacttc	174060
atcaatcaaa	tatttttttga	gtgtatgctg	tagctgaaag	agtatgtacg	tacgtataag	174120
actagagaga	tattaagtct	cagtacactt	cctgtgccat	gttattcagc	tcactggttt	174180
acaaatatag	gttgtcttgt	ggttgttagga	gccactgta	acaatactgg	gcagcctttt	174240
tttttttttt	tttaattgca	acaatgcaaa	agccaagaaa	gtataagggt	cacaagtcta	174300
aacaatgaat	tcnttcaaca	gggaaaacag	ctagcttgaa	aacttgctga	aaaacacaac	174360
ttgtgtttat	ggcatttagt	accttcaaat	aattggcctt	gcagatattg	gataccccat	174420
taaatctgac	agtctcaaat	ttttcatctc	ttcaatcact	agtcaagaaa	aatataaaaa	174480
caacaaatac	ttccatatgg	agcatttttc	agagttttct	aaccagctct	tattttttcta	174540
gtcagtaaac	atttgtaaaa	atactgtttc	actaatactt	actgttaact	gtcttgagag	174600
aaaagaaaaa	tatgagagaa	ctattgtttg	gggaagttca	agtgatcttt	caatatcatt	174660
actaacttct	tccacttttt	ccagaatttg	aatattaacg	ctaaagggtg	aagacttcag	174720
atttcaaat	aatcttttcta	tattttttta	atttacagaa	tattatataa	cccactgctg	174780
aaaaagaaaa	aaatgattgt	tttagaagtt	aaagtcaata	ttgattttta	atataagtaa	174840
tgaaggcata	tttccaataa	ctagtgatat	ggcatcggtg	catttttacag	tatcttcaaa	174900
aatacagaat	ttatagaata	atttctcctc	atttaaatatt	tttcaaaatc	aaagttatgg	174960
tttctctcatt	ttactaaaat	cgtattctaa	ttcttcatta	tagtaaatct	atgagcaact	175020
ccttacttctg	gttctctctga	tttcaaggcc	atatttttaaa	aaatcaaaaag	gcactgtgaa	175080
ctattttgaa	gaaaacacaa	catttttaata	cagattgaaa	ggacctcttc	tgaagctaga	175140

aacaatctat	agttatacat	cttcattaat	actgtgttac	cttttaaaat	agtaatTTTT	175200
tacattttcc	tgtgtaaacc	taattgtggt	agaaatTTTT	accaactcta	tactcaatca	175260
agcaaaatTT	ctgtatatTC	cctgtggaat	gtacctatgt	gagtttcaga	aattctcaaa	175320
atacgtgttc	aaaaatTTCT	gcttttgcat	ctttgggaca	cctcagaaaa	cttattaaca	175380
actgtgaata	tgagaaatac	agaagaaaaT	aataagccct	ctatacataa	atgccagca	175440
caatttcattg	ttaaaaaaca	accaaaccctc	acactactgt	atttcattat	ctgtactgaa	175500
agcaaatgct	ttgtgactat	taaatgttgc	acatcattca	ttcactgtat	agtaatcatt	175560
gactaaagcc	atttgtctgt	gttttcttct	tgtggttgta	tatatcaggt	aaaatatttt	175620
ccaaagagcc	atgtgtcatg	taatactgaa	ccactttgat	attgagacat	taatttgtac	175680
ccttgttatt	atctactagt	aataatgtaa	tactgtagaa	atattgctct	aattcttttc	175740
aaaattgttg	catccccctt	agaatgtttc	tatttccata	aggatttagg	tatgctatta	175800
tcccttctta	taccctaaga	tgaagctggt	tktgtgctct	ttgttcatca	ttggccctca	175860
ttccaagcac	tttacgctgt	ctgtaatgsg	atctattttt	gcactggaat	atctgagaaa	175920
ttgmamaact	agacaaaagt	ttcacaaacag	atttctaagt	taaatcattt	tcatwaaarr	175980
raaaaararaa	aaaaaatttt	gtatgtcaat	aacttttatat	gaagtattaa	aatgcatatt	176040
tctatgttgt	aatataatga	gtcacaaaat	aaagctgtga	cagttctggt	gggtctacaga	176100
aattttacttt	tgtgcatttg	tggcaccacc	tactgttgaa	gggttataaa	gccattagaa	176160
aagtagaggg	gaagtgattt	ggatcaaaaag	gaaaaacttt	agaaaagatt	caaagtgtcc	176220
cttaatcata	aaagagaact	gaggggacta	cttgaaaata	aaagggttgtt	ttgtattttc	176280
atgtttggtta	agatactgag	ntaacwggta	ttaagtgtta	gaggttttta	gataaatatt	176340
ctgcttaatg	attatgaagc	tgcactgaga	tttctgaaaa	tgctctgtag	ctgagcttat	176400
ttaataaaatg	ttcacttggt	ataggggaag	ctacaaaggc	agccttcagt	gtccttttgt	176460
ttattcaacc	aaaaatataa	ggacacaatg	tagcagttat	actgggaagg	tgctgggggt	176520
ggtggcaatg	gtgagcagga	aggcgaagta	gatatggaaa	cagaaatgat	actaatatcg	176580
gtgattccct	ccttttttcc	tgtrataagt	gctgtgcaga	caacatatga	gcagtgtctga	176640
taaatgtaaa	tgtattkttc	atagctcatt	aagaatcagt	ttcagaaaga	gatgtctgct	176700
tatttkgctr	cctgaagaat	ccctgtcaaa	cagtcccttt	saggaagtac	aagaggctgt	176760
ctctatttgt	gacctcagga	atggctgtga	cagtgtcgtg	agcagtcctt	ttcctgtggc	176820
acagatctga	actttgtgtg	cagaaaaatc	ttggcttcaa	gtgagccaag	atgccccctg	176880
agcatcagca	tcacaacttc	atcctcctat	cttgaagttc	atgttatagt	gactttaatg	176940
aaatcataga	acactgtttc	ttcgtgnaac	aatgacgagg	gagaggaaaa	aactttattg	177000
aaaaataaaa	aggcaggtaa	tttagatgaa	aatatgttac	ccatgagggt	ttgtttttgc	177060
tttttgTTTT	tgTTTTtgag	aaacagaatc	tcgctctgtc	gtccaggctg	gagtgcagcg	177120
gcatgatctt	ggctcactgc	aacctccgcc	tcccgggttc	aagcgattct	cctcagcttc	177180
ccaagtagct	ggtactacag	gcatgcgcc	ccacaaccag	ctaatttttg	tatttttagt	177240
agagatgggg	tttactata	cgttggccag	gctggtctca	aactcctgac	ctaagggtgat	177300
ccttctgcct	tgggctccca	aagtgtggg	attacaggca	tgagccacct	tgctggccc	177360
tacccatgag	ccttgactaa	aacattcttc	tatctgtaga	aaagcccaaa	agaacttttc	177420
cagattcaaa	aaacttgga	ctttgtaatg	gtaatgttta	cattaagtaa	aaaaaaaaaa	177480
aaaacccact	tagcttcagt	tttcaagtgt	ttactgtgtt	gtcatgcaact	tcattttaatt	177540
ctcaacacct	gccctatgag	gtaaaaagta	ccattttaca	tatgagtaaa	ttacagctca	177600
gtggataaga	aactcgtcca	aaggtacagg	ttcagtcagg	tggcagagg	ttctttttgt	177660
tgaagttagg	tatcagttaa	aattgacctt	gtaaaatcac	atcagcatca	atatacatTA	177720
atttaacaaa	tatttattga	actttactgt	atgccagata	cttctctagg	tactaggggg	177780
tacaatgtag	aagaaaatag	aattcctgct	ctaatagaatt	tatatttcag	tggtgaaaga	177840
tgatgtgtgg	acaaacacat	ctaatagtatt	ttgacagcaa	tgggtgctaa	gaagaaaata	177900
agacatggta	attggataac	aatggagaga	agtcagagat	ggccttttct	gaggagtgc	177960
atgctcagaa	tcgaataaaa	agagcaggag	cagcctttca	aggtgggtgc	gaaggatatt	178020
cccgaagaa	agaataagtg	ttgcaaaagc	cccagtggga	aaaagcttgg	caagctgagg	178080
tggtaaaaga	gctatctgta	ctaatagtctc	ttgtactgag	ttagcaagga	caagagtggc	178140
aggggggtgga	actggagaga	tagcaggggac	cacatctcac	aggatctcgc	cagccttttt	178200
aagggtatttg	gattttattg	ttagtgcac	aaggagccac	tggagagttt	taaacagtag	178260
tggtgtgacc	tactgtttca	gaaagaacac	tggctactat	gggaagaaag	gacagtagga	178320
agactaatag	ataatggtgg	attgaactaa	gatggttagca	acagatagg	agctatggta	178380

atgttcagta	tccacttttg	agatagatcc	agcaggactt	gctgacagac	tgcatgtaag	178440
ggctgagga	aagagagcta	gcaagggtgac	ttctagtttg	tgacctgaac	taggtagatg	178500
gttgtgttaa	aaatgcagtg	ggtattttcag	tctgaagctc	aggggcaatg	atggaaaata	178560
agataaatgt	gtgcaagttg	gctgggcgtg	gtggtgcacg	cctataatca	cagcactttg	178620
ggaggccgag	acgggtggat	cacttgagggt	caggcggttg	cgaccagcct	aacatggtga	178680
aaccccatct	ctaataaata	caaaaaaatt	agctgggcat	ggtggtgcat	gcctgtaatc	178740
caagctactt	gggaggctga	gacaagagaa	tcgcttgac	ctgggaggcg	gaggtttcag	178800
tgagctgaga	tcgcaccatt	gcactccagc	ctgggcaaca	agagcgaaac	tccatctaaa	178860
aaaaaaagt	tgcaagttat	tggtacgtat	atatttaagc	ctcgggactg	aatgagatta	178920
cctagagaga	atgaagatag	ggaacattct	tgagcctgag	atcctataac	atttagaaga	178980
gaagctagcc	agggaaattg	agagggactg	gccagtgaga	gagagtaaaa	tctgaagagt	179040
ataatgccat	atgttgaagc	tcagagaatt	attgtttcag	gaaggcagaa	gttgtcaatt	179100
atgtccaaca	ttgctgagaa	gatgagtaag	gaatgtaa	gagtaaggc	aggaatggct	179160
attgaatttg	ttaagacaga	ggtccttgggt	gaccttccta	agagtcattt	cagcatgaca	179220
aggatgcaga	cctaactgga	atagattttc	aggggaagact	gagaggcaag	gaaggggaga	179280
ctgcaaatat	atgcaaactc	atatttttgc	tttgaaggga	aacagaaaac	aggacaacag	179340
ctgggtttaca	gaacttttat	ttggattcct	tccttgtctg	ttaataatat	aggctgcacc	179400
tgggcatcat	cgtgcccttc	ctttttgatt	ctgctgtttt	gcttttaatc	ctacaggcat	179460
ttcttctggc	aggagccttg	tacacactct	gggtggttca	tgaccaccac	ggacttgccc	179520
tttcaaggca	gaagcctgtg	attcaaaaac	ttccattcct	tccaatcgac	tattataaaa	179580
ctcattttgt	ctagtgccta	tactccttta	ccttaggatc	ctgggccaac	tgtcactaaa	179640
gaaaacaaac	catttctctc	tcctggggga	atgtgcttct	ttcaagaagg	tgcttagaga	179700
atagacattc	taggtgacag	cagttatgaa	agaggtgtct	ctaatagaacc	taacacttct	179760
tccaacatac	tccaggttca	ccctcctttg	aagacagaaa	gtactacaac	ttcagttgcc	179820
accctagtca	ggttcatagg	atttactcag	taggttaaga	gatggtgctg	agaatgaagt	179880
ccagagtttg	agtctacaag	tggcatttgc	ttctcctttg	ggacttctgt	ggcaggggtg	179940
ggaaccaa	cagattgggt	gcatactctg	ttctcttctc	gaaaatacaa	ttcaaatacc	180000
ttcatagtct	gtgacagcaa	cttataacca	acagacaatg	gtattatttt	tatttattga	180060
tggcactccc	ccatagccag	aacttctact	gagtcattca	tcgtgtattc	caggaatact	180120
ttccaaatgg	agtagtgagg	cacttccaat	aacagttatg	taatctgtta	acagtggctc	180180
aaaatgacat	gtgaaggaat	gatgtacttc	ttttttcaga	ctcacagatc	tcatatgtta	180240
attcagtcaa	caagtactta	gggcttatca	catgccaggc	acaggttaata	tgacaataat	180300
gaaaacaaaa	acccttgctg	tcatagaatt	cggtttctag	tgggaagaga	catacattaa	180360
caaatacatat	ataagtaaaa	catttaattat	ggtgataaca	gaaaaagaaa	gcaggggaga	180420
gggatagtgt	cttagttcag	gttgctgtaa	caagagtacc	atagacaggg	tggcttamac	180480
aacaaacatt	tctttctcac	agatctagtg	gctgggaagt	ctaagaccag	ggtgccagca	180540
tgaagaagtt	ctggtgaagg	cttacattct	ggtttcctca	cagtgagggc	tcttcacgcc	180600
cttataaatg	cactaattct	attaccacac	tgacctcatc	taaaactaac	tcacaaaggc	180660
caaacctcct	atgttgagat	taaggcttca	acataggaat	tttgggggga	caaaaacatt	180720
cagtctctag	cagatagggg	ccactgtggg	aggggattac	aatcttaaat	agggttgtca	180780
gggaaggcct	catgagggca	aagacctgag	gaggcatgga	aatgatccac	gcagatgcat	180840
gtctggggaa	gagcattcca	ggcagaggca	gtgggaatgg	accctcaggt	aagagcatgt	180900
ttagcaagtt	cagagaaaag	cctcttaggt	caggtctgga	gcgagagtgg	aaaatgtaac	180960
aggacagatc	atgtaaagac	ttgtagacca	cagttagagg	cttggacaat	aactctgagg	181020
atgacatgaa	ggtattggga	gggggtgttg	agcagaggag	tgatatgac	tggtttata	181080
tttagattat	tctgactgct	atgttgaaaa	tagggttgag	agcagagggg	aagggcagca	181140
caccagttaa	gaccacttcc	caaaaatcca	gatgatgatg	ggcttgaacc	aggagagttg	181200
ctatagggat	gtgaagcaat	caaattctag	atatttggga	cctgctgaca	ggtacataat	181260
ggtgtatgag	agaaaaggca	gtgtgagaat	gactctgaag	tctgtggcct	gagcaaaaaca	181320
gaaacctgga	gttgccaata	gctgagataa	agtttctggt	taaagatcaa	gagctcagtt	181380
ttggacaaga	ggggtgtgag	atttctaata	gacattcagg	tggaggtgtg	gggaacacac	181440
ctggatgtga	tgagcctgga	attcaaattc	acatgtgtga	atttgaaaaa	agattcagaa	181500
aagaggtgtg	gactggagga	gctatccaca	cacagatagc	atttgaagcc	acaagactgg	181560
gggtgatgac	gcagggagta	agtataaata	gaaaagagaa	gtggtccaag	gactcagtcc	181620

ttacatcaga	ggtcaaagag	acaagatgga	gactaggaga	tatgaaggct	tatgcagtga	181680
gtgtttcggg	gaaacgaaga	aacttcccag	aagaaaggag	aggccaactg	ctaacaagtc	181740
aagtgaatg	agaactggga	aatgagttga	caagagcaat	ttggctggcg	tgatcagaga	181800
atgtttaaga	gataatgtga	gggaagaaaa	tggaggcagg	aagtatgaac	tactcttttg	181860
agtcataattg	taaaggaaaa	aaatggggtg	atcttactct	tttttttttt	taaagatagg	181920
aaaaaaaaagt	atatttatat	gctgaggaga	agaatccagt	tgaaggggaa	tactgtggct	181980
tcaggagaca	ggaaaactgc	ctaagccaca	tccttcagca	agtgagtgtt	caagtggagg	182040
gtttgtcttt	gcttggggca	cgggcggttc	atccacagtg	aaaggagaga	aggcagggca	182100
tgtgggggttg	gctgtagaaa	gcgggcagac	gggggtgggaa	cttatgaaga	ctcttctgat	182160
tgttattttc	tctgttaagt	aaaaggaaaag	gtcattggct	aagaatgaag	atggacgcgc	182220
agggtgctggg	agtcacgaa	gagaatgaaa	gaatgaataa	aatagagaaa	aatacaatgc	182280
tattatataa	acctatacat	ttaaattact	ttcctcttca	aattagaccc	cctggaaggc	182340
aaatctctgt	catgttaagt	tttaggaaaa	gtcataatct	acttgagagct	aagtatgaag	182400
aataagggtat	atgatttaaa	ccacataaatt	ctatttttgc	ttaaaccggg	gaagccttta	182460
aggtaacgta	gagtgtaccc	attaaaaaag	ataaaaaatgt	cagcttcccc	ttccctttta	182520
atatagtgaag	gctcaaagg	aattgatgtt	cataacccta	aggctactata	acagctctgaa	182580
gcaaaatatt	aaatgtcttc	tgtagtattt	tagttaactt	ttaattttat	tggtaattta	182640
tgtgtagtta	accacatact	tgtgcctcaa	ctatttcaca	taaatgtgag	atgtcttagg	182700
cctctttgaa	ctctgcatgg	aaaagaatgc	tgaggaggta	tgtctgttta	agctgccact	182760
gcctgcctct	cataaataat	atctcacatt	gcataatagc	ccagataaac	tcagatcgc	182820
ttagagcctg	ataaatgaga	ggaaactgta	cagtttgttt	tccaaattgc	ccttgagcag	182880
ctaattcaca	atttactatt	ttatagctgt	aagattttta	aaatataaaa	taacatagct	182940
gtaaatacag	attatcagaa	gatagctgt	aaaaatgtaa	atataatata	gctgtaaata	183000
tacactatcg	gaagtaccaa	atctagtttt	agtaggtaga	ggcatcaagt	aaaatggaaa	183060
tgaagatgt	ttcaaggaaa	actccaggac	atctgtggca	gtactaaaga	aaccttcctt	183120
ctcaggttcc	cagcatgcta	ttttattgat	gtaacaacat	tttcaataaa	gttggtaaat	183180
tatcactata	ttactgtctt	atagcaacat	agcaagagct	ttagagatca	tataagtata	183240
aaatgtgaat	ttaaaaaaac	aatgaatatg	caggattttt	attagggcaa	gcgtttccat	183300
aaccataaat	atttctttta	aacaaataaa	tgtccaaga	tctctgttag	tgatccaaac	183360
taagtagaaa	ttagtataat	taattataaa	tgaacaattt	cagcatataa	accaacaagt	183420
cttttctaga	tttttaacac	tgtgacccaa	ttgcattatt	ttccaagtta	gaatgactaa	183480
taatcaatga	atgtaaaagc	aataattaat	acagatgaca	ttgtactttt	ccacagtaaa	183540
gaaataaaca	atctaataat	tttataaatc	ccattttata	tcacaaaata	acctttacta	183600
agcaaatttt	tttaaaatct	caggaactat	agacatgatg	aaaagataga	tattttatat	183660
aaataattca	aaaatactgt	caggcaagga	aatgtaaaat	ccttatttga	gtaaaagaaa	183720
atgctataaa	gcaatgagtt	atcaaaaatac	agaagaggta	ttctaaaaca	aatgaaaaaac	183780
caagatgatg	aaatagtgaac	aactacttct	aatgtgtaac	agatactgaa	atgccaaagg	183840
gaaagtgaac	tgaattattt	cttaaaagcag	tggagaatat	gtaactttca	aaaatgcaag	183900
aagcacagca	aattaactaa	ttaactttacc	acctcttca	aataaaaagcg	agaacctctt	183960
gggagaattt	aagcaccatt	agcagacaca	tcttagagc			183999

<210> 2
 <211> 21
 <212> DNA
 <213> homo sapien

<400> 2
 ggaaagtgttc taaaagagaa a

21

<210> 3
 <211> 21
 <212> DNA
 <213> homo sapien

<400> 3
 agtaaagagg gactagactt t 21

<210> 4
 <211> 21
 <212> DNA
 <213> homo sapien

<400> 4
 agtaaagagg aactagactt t 21

<210> 5
 <211> 2261
 <212> PRT
 <213> homo sapien

<400> 5
 Met Ala Cys Trp Pro Gln Leu Arg Leu Leu Leu Trp Lys Asn Leu Thr
 1 5 10 15
 Phe Arg Arg Arg Gln Thr Cys Gln Leu Leu Leu Glu Val Ala Trp Pro
 20 25 30
 Leu Phe Ile Phe Leu Ile Leu Ile Ser Val Arg Leu Ser Tyr Pro Pro
 35 40 45
 Tyr Glu Gln His Glu Cys His Phe Pro Asn Lys Ala Met Pro Ser Ala
 50 55 60
 Gly Thr Leu Pro Trp Val Gln Gly Ile Ile Cys Asn Ala Asn Asn Pro
 65 70 75 80
 Cys Phe Arg Tyr Pro Thr Pro Gly Glu Ala Pro Gly Val Val Gly Asn
 85 90 95
 Phe Asn Lys Ser Ile Val Ala Arg Leu Phe Ser Asp Ala Arg Arg Leu
 100 105 110
 Leu Leu Tyr Ser Gln Lys Asp Thr Ser Met Lys Asp Met Arg Lys Val
 115 120 125
 Leu Arg Thr Leu Gln Gln Ile Lys Lys Ser Ser Asn Leu Lys Leu
 130 135 140
 Gln Asp Phe Leu Val Asp Asn Glu Thr Phe Ser Gly Phe Leu Tyr His
 145 150 155 160
 Asn Leu Ser Leu Pro Lys Ser Thr Val Asp Lys Met Leu Arg Ala Asp
 165 170 175
 Val Ile Leu His Lys Val Phe Leu Gln Gly Tyr Gln Leu His Leu Thr
 180 185 190
 Ser Leu Cys Asn Gly Ser Lys Ser Glu Glu Met Ile Gln Leu Gly Asp
 195 200 205
 Gln Glu Val Ser Glu Leu Cys Gly Leu Pro Arg Glu Lys Leu Ala Ala
 210 215 220
 Ala Glu Arg Val Leu Arg Ser Asn Met Asp Ile Leu Lys Pro Ile Leu
 225 230 235 240
 Arg Thr Leu Asn Ser Thr Ser Pro Phe Pro Ser Lys Glu Leu Ala Glu

1105		1110		1115		1120
Gln Leu Gly Thr	Gly Tyr Tyr Leu Thr	Leu Val Lys Lys Asp Val Glu				
	1125		1130		1135	
Ser Ser Leu Ser Ser Cys Arg Asn Ser Ser Ser Thr Val Ser Tyr Leu						
	1140		1145		1150	
Lys Lys Glu Asp Ser Val Ser Gln Ser Ser Ser Asp Ala Gly Leu Gly						
	1155		1160		1165	
Ser Asp His Glu Ser Asp Thr Leu Thr Ile Asp Val Ser Ala Ile Ser						
	1170		1175		1180	
Asn Leu Ile Arg Lys His Val Ser Glu Ala Arg Leu Val Glu Asp Ile						
1185		1190		1195		1200
Gly His Glu Leu Thr Tyr Val Leu Pro Tyr Glu Ala Ala Lys Glu Gly						
	1205		1210		1215	
Ala Phe Val Glu Leu Phe His Glu Ile Asp Asp Arg Leu Ser Asp Leu						
	1220		1225		1230	
Gly Ile Ser Ser Tyr Gly Ile Ser Glu Thr Thr Leu Glu Glu Ile Phe						
	1235		1240		1245	
Leu Lys Val Ala Glu Glu Ser Gly Val Asp Ala Glu Thr Ser Asp Gly						
	1250		1255		1260	
Thr Leu Pro Ala Arg Arg Asn Arg Arg Ala Phe Gly Asp Lys Gln Ser						
1265		1270		1275		1280
Cys Leu Arg Pro Phe Thr Glu Asp Asp Ala Ala Asp Pro Asn Asp Ser						
	1285		1290		1295	
Asp Ile Asp Pro Glu Ser Arg Glu Thr Asp Leu Leu Ser Gly Met Asp						
	1300		1305		1310	
Gly Lys Gly Ser Tyr Gln Val Lys Gly Trp Lys Leu Thr Gln Gln Gln						
	1315		1320		1325	
Phe Val Ala Leu Leu Trp Lys Arg Leu Leu Ile Ala Arg Arg Ser Arg						
	1330		1335		1340	
Lys Gly Phe Phe Ala Gln Ile Val Leu Pro Ala Val Phe Val Cys Ile						
1345		1350		1355		1360
Ala Leu Val Phe Ser Leu Ile Val Pro Pro Phe Gly Lys Tyr Pro Ser						
	1365		1370		1375	
Leu Glu Leu Gln Pro Trp Met Tyr Asn Glu Gln Tyr Thr Phe Val Ser						
	1380		1385		1390	
Asn Asp Ala Pro Glu Asp Thr Gly Thr Leu Glu Leu Leu Asn Ala Leu						
	1395		1400		1405	
Thr Lys Asp Pro Gly Phe Gly Thr Arg Cys Met Glu Gly Asn Pro Ile						
	1410		1415		1420	
Pro Asp Thr Pro Cys Gln Ala Gly Glu Glu Glu Trp Thr Thr Ala Pro						
1425		1430		1435		1440
Val Pro Gln Thr Ile Met Asp Leu Phe Gln Asn Gly Asn Trp Thr Met						
	1445		1450		1455	
Gln Asn Pro Ser Pro Ala Cys Gln Cys Ser Ser Asp Lys Ile Lys Lys						
	1460		1465		1470	
Met Leu Pro Val Cys Pro Pro Gly Ala Gly Gly Leu Pro Pro Pro Gln						
	1475		1480		1485	
Arg Lys Gln Asn Thr Ala Asp Ile Leu Gln Asp Leu Thr Gly Arg Asn						
	1490		1495		1500	
Ile Ser Asp Tyr Leu Val Lys Thr Tyr Val Gln Ile Ile Ala Lys Ser						
1505		1510		1515		1520
Leu Lys Asn Lys Ile Trp Val Asn Glu Phe Arg Tyr Gly Gly Phe Ser						
	1525		1530		1535	
Leu Gly Val Ser Asn Thr Gln Ala Leu Pro Pro Ser Gln Glu Val Asn						

1540	1545	1550
Asp Ala Ile Lys Gln Met Lys Lys His Leu Lys Leu Ala Lys Asp Ser		
1555	1560	1565
Ser Ala Asp Arg Phe Leu Asn Ser Leu Gly Arg Phe Met Thr Gly Leu		
1570	1575	1580
Asp Thr Arg Asn Asn Val Lys Val Trp Phe Asn Asn Lys Gly Trp His		
1585	1590	1595
Ala Ile Ser Ser Phe Leu Asn Val Ile Asn Asn Ala Ile Leu Arg Ala		
1605	1610	1615
Asn Leu Gln Lys Gly Glu Asn Pro Ser His Tyr Gly Ile Thr Ala Phe		
1620	1625	1630
Asn His Pro Leu Asn Leu Thr Lys Gln Gln Leu Ser Glu Val Ala Leu		
1635	1640	1645
Met Thr Thr Ser Val Asp Val Leu Val Ser Ile Cys Val Ile Phe Ala		
1650	1655	1660
Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu Ile Gln Glu Arg		
1665	1670	1675
Val Ser Lys Ala Lys His Leu Gln Phe Ile Ser Gly Val Lys Pro Val		
1685	1690	1695
Ile Tyr Trp Leu Ser Asn Phe Val Trp Asp Met Cys Asn Tyr Val Val		
1700	1705	1710
Pro Ala Thr Leu Val Ile Ile Ile Phe Ile Cys Phe Gln Gln Lys Ser		
1715	1720	1725
Tyr Val Ser Ser Thr Asn Leu Pro Val Leu Ala Leu Leu Leu Leu		
1730	1735	1740
Tyr Gly Trp Ser Ile Thr Pro Leu Met Tyr Pro Ala Ser Phe Val Phe		
1745	1750	1755
Lys Ile Pro Ser Thr Ala Tyr Val Val Leu Thr Ser Val Asn Leu Phe		
1765	1770	1775
Ile Gly Ile Asn Gly Ser Val Ala Thr Phe Val Leu Glu Leu Phe Thr		
1780	1785	1790
Asp Asn Lys Leu Asn Asn Ile Asn Asp Ile Leu Lys Ser Val Phe Leu		
1795	1800	1805
Ile Phe Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Lys		
1810	1815	1820
Asn Gln Ala Met Ala Asp Ala Leu Glu Arg Phe Gly Glu Asn Arg Phe		
1825	1830	1835
Val Ser Pro Leu Ser Trp Asp Leu Val Gly Arg Asn Leu Phe Ala Met		
1845	1850	1855
Ala Val Glu Gly Val Val Phe Phe Leu Ile Thr Val Leu Ile Gln Tyr		
1860	1865	1870
Arg Phe Phe Ile Arg Pro Arg Pro Val Asn Ala Lys Leu Ser Pro Leu		
1875	1880	1885
Asn Asp Glu Asp Glu Asp Val Arg Arg Glu Arg Gln Arg Ile Leu Asp		
1890	1895	1900
Gly Gly Gly Gln Asn Asp Ile Leu Glu Ile Lys Glu Leu Thr Lys Ile		
1905	1910	1915
Tyr Arg Arg Lys Arg Lys Pro Ala Val Asp Arg Ile Cys Val Gly Ile		
1925	1930	1935
Pro Pro Gly Glu Cys Phe Gly Leu Leu Gly Val Asn Gly Ala Gly Lys		
1940	1945	1950
Ser Ser Thr Phe Lys Met Leu Thr Gly Asp Thr Thr Val Thr Arg Gly		
1955	1960	1965
Asp Ala Phe Leu Asn Lys Asn Ser Ile Leu Ser Asn Ile His Glu Val		

1970	1975	1980
His Gln Asn Met Gly Tyr Cys Pro Gln Phe Asp Ala Ile Thr Glu Leu		
1985	1990	1995
Leu Thr Gly Arg Glu His Val Glu Phe Phe Ala Leu Leu Arg Gly Val		2000
	2005	2010
Pro Glu Lys Glu Val Gly Lys Val Gly Glu Trp Ala Ile Arg Lys Leu		2015
	2020	2025
Gly Leu Val Lys Tyr Gly Glu Lys Tyr Ala Gly Asn Tyr Ser Gly Gly		2030
	2035	2040
Asn Lys Arg Lys Leu Ser Thr Ala Met Ala Leu Ile Gly Gly Pro Pro		2045
	2050	2055
Val Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Lys Ala Arg		2060
2065	2070	2075
Arg Phe Leu Trp Asn Cys Ala Leu Ser Val Val Lys Glu Gly Arg Ser		2080
	2085	2090
Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu Cys Thr		2095
	2100	2105
Arg Met Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu Gly Ser Val		2110
	2115	2120
Gln His Leu Lys Asn Arg Phe Gly Asp Gly Tyr Thr Ile Val Val Arg		2125
	2130	2135
Ile Ala Gly Ser Asn Pro Asp Leu Lys Pro Val Gln Asp Phe Phe Gly		2140
2145	2150	2155
Leu Ala Phe Pro Gly Ser Val Leu Lys Glu Lys His Arg Asn Met Leu		2160
	2165	2170
Gln Tyr Gln Leu Pro Ser Ser Leu Ser Ser Leu Ala Arg Ile Phe Ser		2175
	2180	2185
Ile Leu Ser Gln Ser Lys Lys Arg Leu His Ile Glu Asp Tyr Ser Val		2190
	2195	2200
Ser Gln Thr Thr Leu Asp Gln Val Phe Val Asn Phe Ala Lys Asp Gln		2205
	2210	2215
Ser Asp Asp Asp His Leu Lys Asp Leu Ser Leu His Lys Asn Gln Thr		2220
2225	2230	2235
Val Val Asp Val Ala Val Leu Thr Ser Phe Leu Gln Asp Glu Lys Val		2240
	2245	2250
Lys Glu Ser Tyr Val		2255
	2260	

<210> 6
 <211> 7860
 <212> DNA
 <213> homo sapien

<400> 6
 gtccctgctg tgagctctgg ccgctgcctt ccagggctcc cgagccacac gctgggggtg 60
 ctggctgagg gaacatgggt tgttgccctc agctgaggtt gctgctgtgg aagaacctca 120
 ctttcagaag aagacaaaca tgtcagctgt tactggaagt ggctggcct ctatttatct 180
 tcctgatcct gatctctgtt cggctgagct acccacccta tgaacaacat gaatgccatt 240
 ttccaaataa agccatgccc tctgcaggaa cacttccttg gggttcagggg attatctgta 300
 atgccaacaa ccctgtttc cgttaccga ctctgggga ggctcccga gttgttgga 360
 actttaacaa atccattgtg gctgcctgt tctcagatgc tcggaggctt cttttataca 420
 gccagaaaga caccagcatg aaggacatgc gcaaagttct gagaacatta cagcagatca 480
 agaaatccag ctcaaacttg aagcttcaag atttctggtt ggacaatgaa accttctctg 540

ggttcctgta	tcacaacctc	tctctcccaa	agtctactgt	ggacaagatg	ctgagggctg	600
atgtcattct	ccacaaggta	tttttgcaag	gctaccagtt	acatttgaca	agtctgtgca	660
atggatcaaa	atcagaagag	atgattcaac	ttggtgacca	agaagtttct	gagctttgtg	720
gcctaccaag	ggagaaactg	gctgcagcag	agcgagtact	tcgttccaac	atggacatcc	780
tgaagccaat	cctgagaaca	ctaaactcta	catctccctt	cccagcaag	gagctggctg	840
aagccacaaa	aacattgctg	catagtcttg	ggactctggc	ccaggagctg	ttcagcatga	900
gaagctggag	tgacatgcga	caggaggtga	tgtttctgac	caatgtgaac	agctccagct	960
cctccaccca	aatctaccag	gctgtgtctc	gtattgtctg	cgggcatccc	gagggagggg	1020
ggctgaagat	caagtctctc	aactgggtatg	aggacaacaa	ctacaaagcc	ctctttggag	1080
gcaatggcac	tgaggaagat	gctgaaacct	tctatgacaa	ctctacaact	ccttactgca	1140
atgatttgat	gaagaatttg	gagtctagtc	ctctttcccg	cattatctgg	aaagctctga	1200
agccgctgct	cgttggaag	atcctgtata	cacctgacac	tccagccaca	aggcaggtca	1260
tggctgaggt	gaacaagacc	ttccaggaac	tggctgtgtt	ccatgatctg	gaaggcatgt	1320
gggaggaact	cagccccaag	atctggacct	tcatggagaa	cagccaagaa	atggaccttg	1380
tccggatgct	gttgacagc	agggacaatg	accacttttg	ggaacagcag	ttggatggct	1440
tagattggac	agcccaagac	atcgtggcgt	ttttggccaa	gcacccagag	gatgtccagt	1500
ccagtaatgg	ttctgtgtac	acctggagag	aagctttcaa	cgagactaac	caggcaatcc	1560
ggaccatatc	tcgcttcatg	gagtgtgtca	acctgaacaa	gctagaacct	atagcaacag	1620
aagtctggct	catcaacaag	tccatggagc	tgctggatga	gaggaagttc	tgggctggta	1680
ttgtgttcac	tggaattact	ccaggcagca	ttgagctgcc	ccatcatgtc	aagtacaaga	1740
tccgaatgga	cattgacaat	gtggagagga	caaataaaat	caaggatggg	tactgggacc	1800
ctggtcctcg	agctgacccc	tttgaggaca	tgcggtacgt	ctgggggggc	ttcgccctact	1860
tgcaggatgt	ggtggagcag	gcaatcatca	gggtgctgac	gggcaccgag	aagaaaactg	1920
gtgtctatat	gcaacagatg	ccctatccct	gttacgttga	tgacatcttt	ctgcgggtga	1980
tgagccggtc	aatgccccctc	ttcatgacgc	tggcctggat	ttactcagtg	gctgtgatca	2040
tcaagggcat	cgtgtatgag	aaggaggcac	ggctgaaaga	gaccatgcgg	atcatgggcc	2100
tggacaacag	catcctctgg	tttagctggg	tcattagtag	cctcattcct	cttcttgtga	2160
gcgtggcct	gctagtggtc	atcctgaagt	taggaaacct	gctgccctac	agtatccca	2220
gcgtggtgtt	tgtcttcctg	tccgtgtttg	ctgtggtgac	aatcctgcag	tgcttcctga	2280
ttagcacact	cttctccaga	gccaacctgg	cagcagcctg	tgggggcac	atctacttca	2340
cgtgtacact	gccctacgtc	ctgtgtgtgg	catggcagga	ctacgtgggc	ttcacactca	2400
agatcttcgc	tagcctgctg	tctcctgtgg	cttttgggtt	tggctgtgag	tactttgccc	2460
tttttgagga	gcagggcatt	ggagtgcagt	gggacaacct	gtttgagagt	cctgtggagg	2520
aagatggctt	caatctcacc	acttcgggtc	ccatgatgct	gtttgacacc	ttcctctatg	2580
gggtgatgac	ctggtacatt	gaggctgtct	ttccaggcca	gtacggaatt	cccaggccct	2640
ggtattttcc	ttgcaccaag	tcctactggg	ttggcgagga	aagtgatgag	aagagccacc	2700
ctggttccaa	ccagaagaga	atatcagaaa	tctgcatgga	ggaggaacct	accacttga	2760
agctgggcgt	gtccattcag	aacctggtaa	aagtctaccg	agatgggatg	aagtggtgctg	2820
tcgatggcct	ggcactgaat	ttttatgagg	gccagatcac	ctccttcctg	ggccacaatg	2880
gagcggggaa	gacgaccacc	atgtcaatcc	tgaccgggtt	gttccccccg	acctcgggca	2940
ccgcctacat	cctgggaaaa	gacattcgct	ctgagatgag	caccatccgg	cagaacctgg	3000
gggtctgtcc	ccagcataac	gtgctgtttg	acatgctgac	tgtcgaagaa	cacatctggt	3060
tctatgcccg	cttgaaaggg	ctctctgaga	agcacgtgaa	ggcggagatg	gagcagatgg	3120
ccctggatgt	tggtttgcca	tcaagcaagc	tgaaaagcaa	aacaagccag	ctgtcaggtg	3180
gaatgcagag	aaagctatct	gtggccttgg	cctttgtcgg	gggatctaag	gttgtcattc	3240
tggatgaacc	cacagctggg	gtggaccctt	actcccgcag	gggaatatgg	gagctgctgc	3300
tgaaataaccg	acaaggccgc	accattatcc	tctctacaca	ccacatggat	gaagcggacg	3360
tcctggggga	caggattgcc	atcatctccc	atgggaagct	gtgctgtgtg	ggctcctccc	3420
tgtttctgaa	gaaccagctg	ggaacaggct	actacctgac	cttgggtcaag	aaagatgtgg	3480
aatcctccct	cagttcctgc	agaaacagta	gtagcactgt	gtcatacctg	aaaaaggagg	3540
acagtgtttc	tcagagcagt	tctgatgctg	gcctgggcag	cgaccatgag	agtgacacgc	3600
tgaccatcga	tgtctctgct	atctccaacc	tcatcaggaa	gcatgtgtct	gaagcccggc	3660
tggtggaaga	catagggcat	gagctgacct	atgtgctgcc	atatgaagct	gctaaggagg	3720
gagcctttgt	ggaactcttt	catgagattg	atgaccggct	ctcagacctg	ggcattttcta	3780

gttatggcat	ctcagagacg	accctggaag	aaatattcct	caaggtggcc	gaagagagtg	3840
gggtggatgc	tgagacctca	gatggtacct	tgccagcaag	acgaaacagg	cgggccttcg	3900
gggacaagca	gagctgtctt	cgcccgttca	ctgaagatga	tgctgctgat	ccaaagtatt	3960
ctgacataga	cccagaatcc	agagagacag	acttgctcag	tgggatggat	ggcaaagggg	4020
cctaccaggt	gaaaggctgg	aaacttacac	agcaacagtt	tgtggccctt	ttgtggaaga	4080
gactgctaata	tgccagacgg	agtcggaaag	gattttttgc	tcagattgtc	ttgccagctg	4140
tgtttgtctg	cattgccctt	gtgttcagcc	tgatcgtgcc	accctttggc	aagtaccca	4200
gcctggaact	tcagccctgg	atgtacaacg	aacagtacac	atttgtcagc	aatgatgctc	4260
ctgaggacac	gggaacctcg	gaactcttaa	acgccctcac	caaagaccct	ggcttcggga	4320
cccgtgttat	ggaaggaaac	ccaatcccag	acacgccctg	ccaggcaggg	gaggaagagt	4380
ggaccactgc	cccagttccc	cagaccatca	tggacctctt	ccagaatggg	aactggacaa	4440
tgcagaaccc	ttcacctgca	tgccagtgtg	gcagcgacaa	aatcaagaag	atgctgcctg	4500
tgtgtccccc	aggggcaggg	gggctgcctc	ctccacaaag	aaaacaaaac	actgcagata	4560
tccttcagga	cctgacagga	agaaacattt	cggattatct	ggtgaagacg	tatgtgcaga	4620
tcatagccaa	aagcttaaa	aacaagatct	gggtgaatga	gttttaggtat	ggcggctttt	4680
ccctgggtgt	cagtaatact	caagcacttc	ctccgagtca	agaagttaat	gatgccatca	4740
aacaaatgaa	gaaacaccta	aagctggcca	aggacagttc	tgcagatcga	tttctcaaca	4800
gcttgggaag	atztatgaca	ggactggaca	ccagaaataa	tgtcaagggtg	tggttcaata	4860
acaagggctg	gcatgcaatc	agctctttcc	tgaatgtcat	caacaatgcc	attctccggg	4920
ccaacctgca	aaaggagag	aacctagcc	attatggaat	tactgctttc	aatcatcccc	4980
tgaatctcac	caagcagcag	ctctcagagg	tggctctgat	gaccacatca	gtggatgtcc	5040
ttgtgtccat	ctgtgtcatc	tttgcaatgt	ccttcgtccc	agccagcttt	gtcgtattcc	5100
tgatccagga	gcgggtcagc	aaagcaaaac	acctgcagtt	catcagtggg	ctgaagcctg	5160
tcacttactg	gctctctaata	tttgtctggg	atatgtgcaa	ttacgttgtc	cctgccacac	5220
tggtcattat	catcttcac	tgcttcagc	agaagtccta	tgtgtcctcc	accaatctgc	5280
ctgtgctagc	ccttctactt	ttgctgtatg	ggtggtcaat	cacacctctc	atgtaccag	5340
cctcctttgt	gttcaagatc	cccagcacag	cctatgtggt	gctcaccagc	gtgaacctct	5400
tcattggcat	taatggcagc	gtggccacct	ttgtgctgga	gctgttcacc	gacaataagc	5460
tgaataatat	caatgatatc	ctgaagtccg	tgttcttgat	cttcccat	ttttgcctgg	5520
gacgagggct	catcgacatg	gtgaaaaacc	aggcaatggc	tgatgccctg	gaaaggtttg	5580
gggagaatcg	ctttgtgtca	ccattatctt	gggacttggt	gggacgaaac	ctcttcgcca	5640
tggccgtgga	aggggtgggtg	ttcttctca	ttactgttct	gatccagtac	agattcttca	5700
tcaggcccag	acctgtaaat	gcaaagctat	ctcctctgaa	tgatgaagat	gaagatgtga	5760
ggcgggaaag	acagagaatt	cttgatgggtg	gaggccagaa	tgacatctta	gaaatcaagg	5820
agttgacgaa	gatatataga	aggaagcgga	agcctgctgt	tgacaggatt	tgcggtgggca	5880
ttcctcctgg	tgagtgcctt	gggctcctgg	gagttaatgg	ggctggaaaa	tcatcaactt	5940
tcaagatggt	aacaggagat	accactgtta	ccagaggaga	tgctttcctt	aacaaaaata	6000
gtatcttata	aaacatccat	gaagtacatc	agaacatggg	ctactgccct	cagtttgatg	6060
ccatcacaga	gctgttgact	gggagagaac	acgtggagtt	ctttgccctt	ttgagaggag	6120
tcccagagaa	agaagtggc	aagggtgggtg	agtgggcgat	tcggaaactg	ggcctcgtga	6180
agtatggaga	aaaatatgct	ggtaactata	gtggaggcaa	caaacgcaag	ctctctacag	6240
ccatggcttt	gatcggcggtg	cctcctgtgg	tgtttctgga	tgaaccacc	acaggcatgg	6300
atcccaaagc	ccggcggttc	ttgtggaatt	gtgccctaag	tgttgtcaag	gaggggagat	6360
cagtagtgct	tacatctcat	agtatggaag	aatgtgaagc	tctttgcact	aggatggcaa	6420
tcatgggtcaa	tggaagggttc	aggtgccttg	gcagtgtcca	gcatctaaaa	aataggtttg	6480
gagatgggtta	tacaatagtt	gtacgaatag	cagggtccaa	cccggacctg	aagcctgtcc	6540
aggattttctt	tggacttgca	tttcttgga	gtgttctaaa	agagaaacac	cggaacatgc	6600
tacaatacca	gcttccatct	tcattatctt	ctctggccag	gatattcagc	atcctctccc	6660
agagcaaaaa	gcgactccac	atagaagact	actctgtttc	tcagacaaca	cttgaccaag	6720
tatttgtgaa	ctttgccaa	gaccaaagtg	atgatgacca	cttaaaagac	ctctcattac	6780
acaaaaacca	gacagtagtg	gacgttgacg	ttctcacatc	ttttctacag	gatgagaaag	6840
tgaaagaaag	ctatgtatga	agaatcctgt	tcatacgggg	tggctgaaag	taaagaggaa	6900
ctagactttc	ctttgcacca	tgtgaagtgt	tgtggagaaa	agagccagaa	gttgatgtgg	6960
gaagaagtaa	actggatact	gtactgatac	tattcaatgc	aatgcaattc	aatgcaatga	7020

aaacaaaatt	ccattacagg	ggcagtgcct	ttgtagccta	tgtcttgtat	ggctctcaag	7080
tgaagactt	gaatttagtt	ttttacctat	acctatgtga	aactctatta	tggaacccaa	7140
tggacatat	ggtttgaact	cacacttttt	tttttttttt	tgttcctgtg	tattctcatt	7200
ggggttgcaa	caataattca	tcaagtaatc	atggccagcg	attattgatc	aaaatcaaaa	7260
ggtaatgcac	atcctcattc	actaagccat	gccatgocca	ggagactggg	ttcccgggtga	7320
cacatccatt	gctggcaatg	agtgtgccag	agttattagt	gccaagtttt	tcagaaagtt	7380
tgaagcacca	tggtgtgtca	tgctcacttt	tgtgaaagct	gctctgctca	gagtctatca	7440
acattgaata	tcagttgaca	gaatgggtgcc	atgcggtggc	aacatcctgc	tttgattccc	7500
tctgataagc	tgttctgggtg	gcagtaacat	gcaacaaaaa	tgtgggtgtc	tccaggcacg	7560
ggaaacttgg	ttccattgtt	atattgtcct	atgcttcgag	ccatgggtct	acagggtcat	7620
ccttatgaga	ctcttaaata	tacttagatc	ctggtaagag	gcaaagaatc	aacagccaaa	7680
ctgctggggc	tgcaactgct	gaagccaggg	catgggatta	aagagattgt	gcgttcaaac	7740
ctaggggaagc	ctgtgccc	ttgtcctgac	tgtctgctaa	catggtacac	tgcatctcaa	7800
gatgtttatc	tgacacaagt	gtattatttc	tggctttttg	aattaatcta	gaaaatgaaa	7860

<210> 7
 <211> 16
 <212> DNA
 <213> homo sapien

<220>
 <221> variation
 <222> (7)...(11)
 <223> N at positions 7-11 is A,T,C, or G

<400> 7
 aggtcannnn aggtca 16

<210> 8
 <211> 26
 <212> DNA
 <213> homo sapien

<400> 8
 agaggcaggt ggatcatttg aggtca 26

<210> 9
 <211> 26
 <212> DNA
 <213> homo sapien

<400> 9
 ttgaggcggg tgatcacttg aggtca 26

<210> 10
 <211> 26
 <212> DNA
 <213> homo sapiens

<400> 10
caaggcgggc agatcacttg aggtta 26

<210> 11
<211> 26
<212> DNA
<213> Homo sapien

<400> 11
caagggtgggc agtcacctc aggtca 26

<210> 12
<211> 24
<212> DNA
<213> Homo sapiens

<220>
<221> variation
<222> (1)...(18)
<223> N at positions 1-6,8,9, and 17 is A,T,C, or G. N
at position 18 is A,T,C, G, or other, including no
nucleotide.

<400> 12
nnnnnnannt tgacctnntg acct 24

<210> 13
<211> 24
<212> DNA
<213> Homo sapiens

<400> 13
ctttgaagcc tgatcatatg acct 24

<210> 14
<211> 24
<212> DNA
<213> Homo sapiens

<400> 14
aggctggtct cgaactcctg acct 24

<210> 15
<211> 24
<212> DNA
<213> Homo sapien

<400> 15
cttaattggt ggwgctgttg acct 24

<210> 16
<211> 24
<212> DNA
<213> Homo sapien

<400> 16	
caggatggcg taaactcctg acct	24
<210> 17	
<211> 24	
<212> DNA	
<213> homo sapien	
<400> 17	
aggttggttt cgaactcctg acct	24
<210> 18	
<211> 24	
<212> DNA	
<213> Homo sapein	
<400> 18	
tcaaggtagg agaccttggt gcct	24
<210> 19	
<211> 10	
<212> DNA	
<213> homo sapien	
<400> 19	
atcacccac	10
<210> 20	
<211> 20	
<212> DNA	
<213> homo sapien	
<400> 20	
gagatgtgct atgacccac	20
<210> 21	
<211> 20	
<212> DNA	
<213> homo sapien	
<400> 21	
gtgagcccag atcacaccac	20
<210> 22	
<211> 20	
<212> DNA	
<213> homo sapien	
<400> 22	
tccatccatc cacacccac	20
<210> 23	
<211> 20	
<212> DNA	

<213> homo sapien

<400> 23

ccctttttatt aacacctcac

20

<210> 24

<211> 20

<212> DNA

<213> homo sapien

<400> 24

gtaagccaag atcatgccac

20

<210> 25

<211> 20

<212> DNA

<213> homo sapien

<400> 25

acctcaagtg atcacccgcc

20

<210> 26

<211> 20

<212> DNA

<213> homo sapien

<400> 26

ggctcaagcg atcctcccac

20

<210> 27

<211> 20

<212> DNA

<213> homo sapien

<400> 27

ccatgattgg atcactgcac

20

<210> 28

<211> 20

<212> DNA

<213> homo sapien

<400> 28

gtgagtcgag atcatgccac

20

<210> 29

<211> 20

<212> DNA

<213> homo sapien

<400> 29

tgcttttggt ttccccccac

20

<210> 30
<211> 20
<212> DNA
<213> homo sapien

<400> 30
ccgccttccc ctcaccccag

20

<210> 31
<211> 20
<212> DNA
<213> homo sapien

<400> 31
accctccacc cccaccccac

20

<210> 32
<211> 16
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (1)...(11)
<223> N at positions 2-9 is A,T,C, or G. N at positions
1 and 10 is A or T. N at position 11 is A or G.

<400> 32
nnnnnnnnnn nggtca

16

<210> 33
<211> 18
<212> DNA
<213> homo sapien

<400> 33
ctgggcaagg atgggtca

18

<210> 34
<211> 17
<212> DNA
<213> homo sapiens

<400> 34
tgggcaagga tgggtca

17

<210> 35
<211> 17
<212> DNA
<213> homo sapiens

<400> 35
aaaaagcacc aaggtca

17

<210> 36
<211> 17
<212> DNA
<213> homo sapien

<400> 36
agaagatgcc aggttca

17

<210> 37
<211> 17
<212> DNA
<213> homo sapiens

<400> 37
gaggagatgg aggttca

17

<210> 38
<211> 26
<212> DNA
<213> homo sapiens

<400> 38
ccgagcgcag aggttactat cgttca

26

<210> 39
<211> 26
<212> DNA
<213> homo sapien

<400> 39
gccattccc aggtcagaac agacca

26

<210> 40
<211> 20
<212> DNA
<213> homo sapien

<400> 40
ggacctgcag ctctccccac

20

<210> 41
<211> 17
<212> DNA
<213> homo sapien

<400> 41
aacgccaag taagtca

17

<210> 42
<211> 17
<212> DNA
<213> homo sapien

<400> 42
gagctcgtac taggaca 17

<210> 43
<211> 17
<212> DNA
<213> homo sapien

<400> 43
gcagagtcct ggggtca 17

<210> 44
<211> 18
<212> DNA
<213> homo sapien

<400> 44
cgcagagtcc tgggggtca 18

<210> 45
<211> 17
<212> DNA
<213> homo sapien

<400> 45
agccaattcc caggtca 17

<210> 46
<211> 17
<212> DNA
<213> homo sapien

<400> 46
acggaccgtt tgggaca 17

<210> 47
<211> 18
<212> DNA
<213> homo sapien

<400> 47
cacggaccgt ttgggaca 18

<210> 48
<211> 19
<212> DNA
<213> homo sapien

<400> 48
ccacggaccg tttgggaca 19

<210> 49
<211> 18
<212> DNA
<213> homo sapien

<400> 49
actagaggcc ttgggtct 18

<210> 50
<211> 17
<212> DNA
<213> homo sapien

<400> 50
ctagaggcct tgggtct 17

<210> 51
<211> 17
<212> DNA
<213> homo sapien

<400> 51
ccctaccct caggtca 17

<210> 52
<211> 18
<212> DNA
<213> homo sapien

<400> 52
tcctacccc tcaggtca 18

<210> 53
<211> 17
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (2)...(2)
<223> N at position 2 is C or G.

<400> 53
gntctgcgcc agggaca 17

<210> 54
<211> 17
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (12)...(12)

<223> N at position 12 is A,T,C, or G.

<400> 54
ttttagtgag anggtta 17

<210> 55
<211> 26
<212> DNA
<213> homo sapien

<400> 55
tgaggcaggt agatcacttg aggtca 26

<210> 56
<211> 26
<212> DNA
<213> homo sapien

<400> 56
cgaggctggc ggatcacctg aggtca 26

<210> 57
<211> 26
<212> DNA
<213> homo sapien

<400> 57
aagcctaaca aggttactga aggcca 26

<210> 58
<211> 26
<212> DNA
<213> homo sapien

<400> 58
agaggtgggc ggatcacctg aggtca 26

<210> 59
<211> 24
<212> DNA
<213> homo sapien

<400> 59
ctcgatttcc tgacctcgtg atcc 24

<210> 60
<211> 24
<212> DNA
<213> homo sapien

<400> 60
caaaacattg tgcccttttg aact 24

<210> 61

<211> 24
<212> DNA
<213> homo sapien

<400> 61
gcgctagggt tgccttcatt acct

24

<210> 62
<211> 24
<212> DNA
<213> homo sapien

<400> 62
ctcgatttct tgacctcgtg atcc

24

<210> 63
<211> 20
<212> DNA
<213> homo sapien

<400> 63
gtgagctgag atcacaccac

20

<210> 64
<211> 20
<212> DNA
<213> homo sapien

<400> 64
ttcaaggatg atcaccacat

20

<210> 65
<211> 20
<212> DNA
<213> homo sapien

<400> 65
ggctcaagtg atcctcccac

20

<210> 66
<211> 20
<212> DNA
<213> homo sapien

<400> 66
gtgagccgag atcgcgccac

20

<210> 67
<211> 20
<212> DNA
<213> homo sapien

<400> 67
gtgagttatg atcatgccac

20

<210> 68
<211> 20
<212> DNA
<213> homo sapien

<400> 68
ccactgtttg aacaaccac

20

<210> 69
<211> 20
<212> DNA
<213> homo sapien

<400> 69
acctcaggtg atccgccac

20

<210> 70
<211> 20
<212> DNA
<213> homo sapien

<400> 70
aaatgtgaca atctccacac

20

<210> 71
<211> 20
<212> DNA
<213> homo sapien

<400> 71
aatatagaat atcacctccc

20

<210> 72
<211> 20
<212> DNA
<213> homo sapien

<400> 72
ccttttatct accaccaac

20

<210> 73
<211> 17
<212> DNA
<213> homo sapien

<400> 73
ccttggtgga tgggtca

17

<210> 74
<211> 18
<212> DNA
<213> homo sapien

<400> 74

gccttggttg atgggtca

18

<210> 75

<211> 17

<212> DNA

<213> homo sapien

<400> 75

ttgctgtgag tgggtca

17

<210> 76

<211> 17

<212> DNA

<213> homo sapien

<400> 76

gccttcgaga aggggtca

17

<210> 77

<211> 18

<212> DNA

<213> homo sapien

<400> 77

ggccttgag aagggtca

18

<210> 78

<211> 18

<212> DNA

<213> homo sapien

<400> 78

aattaagctg atgggtca

18

<210> 79

<211> 17

<212> DNA

<213> homo sapien

<400> 79

attaagctga tgggtca

17

<210> 80

<211> 18

<212> DNA

<213> homo sapien

<400> 80

aggtgctaac taggggtca

18

<210> 81

<211> 17

<212> DNA

<213> homo sapien

<400> 81 ggtgctaact aggggtca	17
<210> 82 <211> 17 <212> DNA <213> homo sapien	
<400> 82 atgggatgac tgggtca	17
<210> 83 <211> 17 <212> DNA <213> homo sapien	
<400> 83 tctccatgcc aagggtca	17
<210> 84 <211> 23 <212> DNA <213> homo sapien	
<400> 84 ttgaggacat gcggtacgtc tgg	23
<210> 85 <211> 23 <212> DNA <213> homo sapien	
<400> 85 ttgaggacat gtggtacgtc tgg	23
<210> 86 <211> 21 <212> DNA <213> homo sapien	
<400> 86 gcctacttgc aggatgtggt g	21
<210> 87 <211> 21 <212> DNA <213> homo sapien	
<400> 87 gcctacttgc gggatgtggt g	21
<210> 88 <211> 23 <212> DNA	

<213> homo sapien

<400> 88

cctcattcct cttcttgtga gcg

23

<210> 89

<211> 20

<212> DNA

<213> homo sapien

<400> 89

cctcattcct cttgtgagcg

20

<210> 90

<211> 21

<212> DNA

<213> homo sapien

<400> 90

aaaagtctac cgagatggga t

21

<210> 91

<211> 21

<212> DNA

<213> homo sapien

<400> 91

aaaagtctac tgagatggga t

21

<210> 92

<211> 21

<212> DNA

<213> homo sapien

<400> 92

ggccagatca cctccttcct g

21

<210> 93

<211> 21

<212> DNA

<213> homo sapien

<400> 93

ggccagatca tctccttcct g

21

<210> 94

<211> 21

<212> DNA

<213> homo sapien

<400> 94

acacaccaca tggatgaagc g

21

<210> 95

<211> 21
 <212> DNA
 <213> homo sapien

 <400> 95
 acacaccaca cggatgaagc g 21

 <210> 96
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 96
 cctggaagaa gtaagttaag t 21

 <210> 97
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 97
 cctggaagaa ctaagttaag t 21

 <210> 98
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 98
 gctgcctgtg tgtccccag g 21

 <210> 99
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 99
 gctgcctgtg cgtccccag g 21

 <210> 100
 <211> 22
 <212> DNA
 <213> homo sapien

 <400> 100
 tagccattat ggaattactg ct 22

 <210> 101
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 101
 tagccattat caattactgc t 21

<210> 102
<211> 26
<212> DNA
<213> homo sapien

<400> 102
gatgaagatg aagatgtgag gcggga

26

<210> 103
<211> 20
<212> DNA
<213> homo sapien

<400> 103
gatgaagatg tgaggcggga

20

<210> 104
<211> 21
<212> DNA
<213> homo sapien

<400> 104
aatagttgta cgaatagcag g

21

<210> 105
<211> 21
<212> DNA
<213> homo sapien

<400> 105
aatagttgta tgaatagcag g

21

<210> 106
<211> 20
<212> DNA
<213> homo sapien

<400> 106
atagttgtac gaatagcagg

20

<210> 107
<211> 19
<212> DNA
<213> homo sapien

<400> 107
atagttgtag aatagcagg

19

<210> 108
<211> 20
<212> DNA
<213> homo sapien

<400> 108 gggtccaacc cggacctgaa	20
<210> 109 <211> 20 <212> DNA <213> homo sapien	
<400> 109 gggtccaacc tggacctgaa	20
<210> 110 <211> 21 <212> DNA <213> homo sapien	
<400> 110 cattatcttc tctggccagg a	21
<210> 111 <211> 20 <212> DNA <213> homo sapien	
<400> 111 cattatcttt ttggccagga	20
<210> 112 <211> 20 <212> DNA <213> homo sapien	
<400> 112 ggaactagtc ccggcaaaaa	20
<210> 113 <211> 20 <212> DNA <213> homo sapien	
<400> 113 ggaactagtc tcggcaaaaa	20
<210> 114 <211> 17 <212> DNA <213> homo sapien	
<400> 114 ccgggacccg cagagcc	17

<210> 115
 <211> 17
 <212> DNA
 <213> homo sapien

 <400> 115
 ccgggaccgg cagagcc 17

 <210> 116
 <211> 20
 <212> DNA
 <213> homo sapien

 <400> 116
 accagccacg gcgtccctgc 20

 <210> 117
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 117
 accagccacg ggcgtccctg c 21

 <210> 118
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 118
 acacgctggg ggtgctggct g 21

 <210> 119
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 119
 acacgctggg cgtgctggct g 21

 <210> 120
 <211> 21
 <212> DNA
 <213> homo sapien

 <400> 120
 ctgggttcct gtatcacaac c 21

 <210> 121
 <211> 21
 <212> DNA
 <213> homo sapien

<400> 121 ctgggttcct atatcacaac c	21
<210> 122 <211> 21 <212> DNA <213> homo sapien	
<400> 122 ggcctaccaa gggagaaact g	21
<210> 123 <211> 21 <212> DNA <213> homo sapien	
<400> 123 ggcctaccaa aggagaaact g	21
<210> 124 <211> 21 <212> DNA <213> homo sapien	
<400> 124 gcgggcatcc cgaggagg g	21
<210> 125 <211> 21 <212> DNA <213> homo sapien	
<400> 125 gcgggcatcc tgaggagg g	21
<210> 126 <211> 21 <212> DNA <213> homo sapien	
<400> 126 agggaggggg gctgaagatc a	21
<210> 127 <211> 21 <212> DNA <213> homo sapien	
<400> 127 agggaggggg actgaagatc a	21

<210> 128
<211> 21
<212> DNA
<213> homo sapien

<400> 128
tgactccagg tgaacaagac c

21

<210> 129
<211> 21
<212> DNA
<213> homo sapien

<400> 129
tgactccagg cgaacaagac c

21

<210> 130
<211> 21
<212> DNA
<213> homo sapien

<400> 130
gcaggactac gtgggcttca c

21

<210> 131
<211> 21
<212> DNA
<213> homo sapien

<400> 131
gcaggactac atgggcttca c

21

<210> 132
<211> 21
<212> DNA
<213> homo sapien

<400> 132
cgtgggcttc acactcaaga t

21

<210> 133
<211> 21
<212> DNA
<213> homo sapien

<400> 133
cgtgggcttc ccactcaaga t

21

<210> 134
<211> 21
<212> DNA
<213> homo sapien

<400> 134 tcacactcaa gatcttcgct g	21
<210> 135 <211> 21 <212> DNA <213> homo sapien	
<400> 135 tcacactcaa catcttcgct g	21
<210> 136 <211> 17 <212> DNA <213> homo sapien	
<400> 136 ccacttcggt ctccatg	17
<210> 137 <211> 17 <212> DNA <213> homo sapien	
<400> 137 ccacttcgat ctccatg	17
<210> 138 <211> 18 <212> DNA <213> homo sapien	
<400> 138 gaagagaata tcagaaag	18
<210> 139 <211> 18 <212> DNA <213> homo sapien	
<400> 139 gaagagaatg tcagaaag	18
<210> 140 <211> 21 <212> DNA <213> homo sapien	
<400> 140 gatctaaggt tgtcattctg g	21

<210> 141
<211> 21
<212> DNA
<213> homo sapien

<400> 141
gatctaaggt ggtcattctg g

21

<210> 142
<211> 20
<212> DNA
<213> homo sapien

<400> 142
gcgaccatga gagtgcacacg

20

<210> 143
<211> 20
<212> DNA
<213> homo sapien

<400> 143
gcgaccatga cagtgacacg

20

<210> 144
<211> 21
<212> DNA
<213> homo sapien

<400> 144
ctggacacca gaaataatgt c

21

<210> 145
<211> 21
<212> DNA
<213> homo sapien

<400> 145
ctggacacca aaaataatgt c

21

<210> 146
<211> 21
<212> DNA
<213> homo sapien

<400> 146
tcctatgtgt cctccaccaa t

21

<210> 147
<211> 21
<212> DNA
<213> homo sapien

<400> 147 tcctatgtgt gctccaccaa t	21
<210> 148 <211> 21 <212> DNA <213> homo sapien	
<400> 148 caggggtccaa cccggacctg a	21
<210> 149 <211> 21 <212> DNA <213> homo sapien	
<400> 149 caggggtccaa tccggacctg a	21
<210> 150 <211> 22 <212> DNA <213> homo sapien	
<400> 150 cgggggaagg gcacgcagac cg	22
<210> 151 <211> 22 <212> DNA <213> homo sapien	
<400> 151 cgggggaagg ggacgcagac cg	22
<210> 152 <211> 17 <212> DNA <213> homo sapien	
<400> 152 ggccgggaac ggggcgg	17
<210> 153 <211> 17 <212> DNA <213> homo sapien	
<400> 153 ggccgggaag ggggcgg	17

<210> 154
<211> 21
<212> DNA
<213> homo sapien

<400> 154
agtatccctt gttcacgaga a 21

<210> 155
<211> 25
<212> DNA
<213> homo sapien

<400> 155
agtatccctc ccttggtcac gagaa 25

<210> 156
<211> 21
<212> DNA
<213> homo sapien

<400> 156
gtgacaccca acggagtagg g 21

<210> 157
<211> 21
<212> DNA
<213> homo sapien

<400> 157
gtgacaccca gcggagtagg g 21

<210> 158
<211> 25
<212> DNA
<213> homo sapien

<400> 158
tatgtgctga ccatgggagc ttgtt 25

<210> 159
<211> 25
<212> DNA
<213> homo sapien

<400> 159
tatgtgctga ccgtgggagc ttgtt 25

<210> 160
<211> 25
<212> DNA
<213> homo sapien

<400> 160
cctccgcctg ccaggttcag cgatt 25

<210> 161
<211> 25
<212> DNA
<213> homo sapien

<400> 161
cctccgcctg ccgggttcag cgatt 25

<210> 162
<211> 21
<212> DNA
<213> homo sapien

<400> 162
gaaaattagt atgtaaggaa g 21

<210> 163
<211> 21
<212> DNA
<213> homo sapien

<400> 163
gaaaattagt ctgtaaggaa g 21

<210> 164
<211> 22
<212> DNA
<213> homo sapien

<400> 164
cattttctta gaaaagagag gt 22

<210> 165
<211> 22
<212> DNA
<213> homo sapien

<400> 165
cattttctta gagaagagag gt 22

<210> 166
<211> 20
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (11)...(11)
<223> N at position 11 is G or T

<400> 166	20
tttaaagggg ntgattagga	
<210> 167	
<211> 22	
<212> DNA	
<213> homo sapien	
<220>	
<221> variation	
<222> (11)...(11)	
<223> N at position 11 is G or T	
<400> 167	22
gaagaaattt ntttttttga tt	
<210> 168	
<211> 19	
<212> DNA	
<213> homo sapien	
<400> 168	19
tctgtcccca tccctgacg	
<210> 169	
<211> 20	
<212> DNA	
<213> homo sapien	
<400> 169	20
tctgtcccca atccctgacg	
<210> 170	
<211> 20	
<212> DNA	
<213> homo sapien	
<400> 170	20
aggagccaaa cgctcattgt	
<210> 171	
<211> 21	
<212> DNA	
<213> homo sapien	
<400> 171	21
aggagccaaa gcgctcattg t	
<210> 172	
<211> 21	
<212> DNA	
<213> homo sapien	

<400> 172
 aagccactgt ttttaaccag t 21

<210> 173
 <211> 21
 <212> DNA
 <213> homo sapien

<400> 173
 aagccactgt atttaaccag t 21

<210> 174
 <211> 22
 <212> DNA
 <213> homo sapien

<400> 174
 gctccctcta gcatgcaggc tc 22

<210> 175
 <211> 22
 <212> DNA
 <213> homo sapien

<400> 175
 gctccctcta gtatgcaggc tc 22

<210> 176
 <211> 21
 <212> DNA
 <213> homo sapien

<400> 176
 ttgcctgttt ctcacagagc c 21

<210> 177
 <211> 19
 <212> DNA
 <213> homo sapien

<400> 177
 ttgcctgttt ctcagagcc 19

<210> 178
 <211> 21
 <212> DNA
 <213> homo sapien

<220>
 <221> variation
 <222> (10)...(10)
 <223> N at position 10 is C or G

<400> 178
gcgagtgcn ctgtgtcctt a 21

<210> 179
<211> 23
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (12)...(12)
<223> N at position 12 is G or T

<400> 179
ctcttctgtt ancacagaag aga 23

<210> 180
<211> 21
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (11)...(11)
<223> N at position 11 is A or G

<400> 180
cattctaggg ntcataagcca t 21

<210> 181
<211> 22
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (11)...(11)
<223> N at position 11 is G or T

<400> 181
aagtacagtg ngaggaacag cg 22

<210> 182
<211> 22
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (12)...(12)
<223> N at position 12 is A or G

<400> 182 attcctaataaa antagaaatg ca	22
<210> 183 <211> 23 <212> DNA <213> homo sapien	
<400> 183 tttctgtttc aattcttgtc tat	23
<210> 184 <211> 23 <212> DNA <213> homo sapien	
<400> 184 tttctgtttc agttcttgtc tat	23
<210> 185 <211> 21 <212> DNA <213> homo sapien	
<400> 185 ggcccctgcc ttattattac t	21
<210> 186 <211> 21 <212> DNA <213> homo sapien	
<400> 186 ggcccctgcc gtattattac t	21
<210> 187 <211> 21 <212> DNA <213> homo sapien	
<400> 187 cactgtctgg gttttaatgt c	21
<210> 188 <211> 21 <212> DNA <213> homo sapien	
<400> 188 cactgtctgg cttttaatgt c	21

<210> 189
<211> 22
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (11)...(11)
<223> N at position 11 is A or G

<400> 189
tgagagaatt ncttgaaccc gg

22

<210> 190
<211> 21
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (11)...(11)
<223> N at position 11 is C or T

<400> 190
tttgctgaaa naatcactga c

21

<210> 191
<211> 22
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (11)...(11)
<223> N at position 11 is C or T

<400> 191
aacctcagtt ncctcatctg tg

22

<210> 192
<211> 21
<212> DNA
<213> homo sapien

<400> 192
aagaagtggc ttgtattttg c

21

<210> 193
<211> 21
<212> DNA
<213> homo sapien

<400> 193
aagaagtggc ctgtattttg c

21

<210> 194
<211> 23
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (11)...(11)
<223> N at position 11 is A or G.

<400> 194
aactgatttg nttggtatag ctg 23

<210> 195
<211> 20
<212> DNA
<213> homo sapien

<400> 195
aataaagata atttcctttgg 20

<210> 196
<211> 20
<212> DNA
<213> homo sapien

<400> 196
aataaagata gtttctttgg 20

<210> 197
<211> 22
<212> DNA
<213> homo sapien

<400> 197
ttcctgcccc gacactcccg cc 22

<210> 198
<211> 22
<212> DNA
<213> homo sapien

<400> 198
ttcctgcccc cacactcccg cc 22

<210> 199
<211> 22
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (11)...(11)
<223> N at position 11 is A or G.

<400> 199 ttgggaggct naggcaggag aa	22
<210> 200 <211> 22 <212> DNA <213> homo sapien	
<400> 200 tgtcagctgt tactggaagt gg	22
<210> 201 <211> 22 <212> DNA <213> homo sapien	
<400> 201 tgtcagctgc tgctggaagt gg	22
<210> 202 <211> 21 <212> DNA <213> homo sapien	
<400> 202 aggagctggc cgaagccaca a	21
<210> 203 <211> 21 <212> DNA <213> homo sapien	
<400> 203 aggagctggc tgaagccaca a	21
<210> 204 <211> 21 <212> DNA <213> homo sapien	
<400> 204 aatgatgcca ccaaacaat g	21
<210> 205 <211> 21 <212> DNA <213> homo sapien	
<400> 205 aatgatgcca tcaaacaat g	21

<210> 206
<211> 21
<212> DNA
<213> homo sapien

<400> 206
gaggtggctc cgatgaccac a

21

<210> 207
<211> 21
<212> DNA
<213> homo sapien

<400> 207
gaggtggctc tgatgaccac a

21

<210> 208
<211> 21
<212> DNA
<213> homo sapien

<400> 208
ttccttaaca gaaatagtat c

21

<210> 209
<211> 21
<212> DNA
<213> homo sapien

<400> 209
ttccttaaca aaaatagtat c

21

<210> 210
<211> 21
<212> DNA
<213> homo sapien

<400> 210
ggaagtgttc caaaagagaa a

21

<210> 211
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 211
gtatttttgc aaggctacca gttacatttg acaa

34

<210> 212
 <211> 28
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 212
 gattggcttc aggatgtcca tgttgga

28

<210> 213
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 213
 gctgctgtga tggggatatct

20

<210> 214
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 214
 acctcactca cacctgggaa

20

<210> 215
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 215
 caagtgagtg cttgggattg

20

<210> 216
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic primer

<400> 216
 tgcttttatt cagggactcc a

21

<210> 217
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 217
gtgatccag cgtggtgttt gtctt

25

<210> 218
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 218
gaaaggccag aggtactcac agcgaagatc ttgaggg

37

<210> 219
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 219
tcgttttatt cagggactcc a

21

<210> 220
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 220
caagtgagtg cttgggattg

20

<210> 221
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 221
cccatgcact gcagagattc

20

<210> 222
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 222
gcaaattcaa atttctccag g

21

<210> 223
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 223
gagaagagcc accctgggtc caaccagaag aggat

35

<210> 224
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 224
aaggcaggag acatcgctt

19

<210> 225
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 225
gagcagttct gatgctggcc tgggcagcga ccacga

36

<210> 226
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 226
tctgcacctc tcctcctctg

20

<210> 227
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 227
cagcttggga agatttatga caggactgga cacga

35

<210> 228
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 228
atgcccctgc caacttac

18

<210> 229
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 229
gtgcaattac gttgtccctg ccacact

27

<210> 230
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic primer

<400> 230
ccatacagca aaagtagaag ggctagcaca

30

<210> 231
<211> 16
<212> DNA
<213> homo sapien

<220>
<221> variation
<222> (7)...(10)
<223> N at position 7-10 is A, T, C, or G

<400> 231
agatcannnn aggtca 16

<210> 232
<211> 16
<212> DNA
<213> Homo sapien

<400> 232
agatcacttg aggtca 16

<210> 233
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 233
cagcgcttcc cgcgcgtctt ag 22

<210> 234
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 234
ccactcactc tcgtccgcaa ttac 24

<210> 235
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 235
ctgctgagtg actgaactac ataaacagag gccgggta 38

<210> 236
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 236
ccactcactc tcgtccgcaa ttac 24

<210> 237
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 237
cagcgcttcc cgcgcgtctt ag 22

<210> 238
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 238
ccactcactc tcgtccgcaa ttac 24

<210> 239
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 239
ctggctttct gctgagtgc 20

<210> 240
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 240
gatcaaagtc cccgaaacc 19

<210> 241
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 241
actcagttgt ataaccact gaaaatgagt 30

<210> 242
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 242
ttctatagat gttatcatct ggg 23

<210> 243
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 243
actcagttgt ataaccact gaaaatgagt 30

<210> 244
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 244
ttctatagat gttatcatct ggg 23

<210> 245
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 245
tcatactaagg cacgttgtgg 20

<210> 246
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 246
cctcaagcct ggagtgactt 20

<210> 247
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 247
atggcaaaca gtcctccaag 20

<210> 248
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 248
accctagcgc tgtgtctctg 20

<210> 249
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 249
atggcaaaca gtcctccaag 20

<210> 250
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 250
accctagcgc tgtgtctctg 20

<210> 251
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 251
tgtgtgtcct cccttcatt 20

<210> 252
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 252
cttggaggac tgtttgccat 20

<210> 253
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 253
cccctcctgc tttatctttc agttaatgac cagccccg 38

<210> 254
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 254
atccccaact caaaaccaca 20

<210> 255
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 255
gccgctgcct tccagggctc ccgagccaca cgctgcg 37

<210> 256
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Primer

<400> 256
atccccaact caaaaccaca 20